RAPID SCREENING OF G PROTEIN-COUPLED RECEPTORS IN YEAST

A Dissertation
Presented to
The Academic Faculty

by

Emily A. Yasi

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry and Biochemistry

Georgia Institute of Technology
August 2020

RAPID SCREENING OF G PROTEIN-COUPLED RECEPTORS IN YEAST

Approved by:

Dr. Pamela Peralta-Yahya, Advisor
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Nicholas Hud
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Amit Reddi
School of Chemistry and Biochemistry
Georgia Institute of Technology

Dr. Mark Styczynski
School of Chemical and Biomolecular Engineering
Georgia Institute of Technology

Dr. Raquel Lieberman
School of Chemistry and Biochemistry
Georgia Institute of Technology

Date Approved: April 21, 2020

For my ‘Dzidzia’ John Piela

ACKNOWLEDGEMENTS

I would not have been able to make it this far without the great people in my life. First, I would like to thank my late grandfather, John Piela, for always encouraging me to continue learning as "education isn't heavy to carry around." I wish I could share this accomplishment with him, but I know he would be proud. I am also grateful for my wonderful parents and first science teachers, Dan and Judy Yasi, who guide and support me to this day. I am beyond grateful to have found my partner, Jessica Dias de Oliveira, who has been a constant calming presence and support throughout this process. I cannot thank her enough for literally everything – cooking, cleaning, and caring for me.

I would also like to thank my high school chemistry and biology teachers Mrs. Pereira and Mrs. Colace, respectively, for giving me a strong foundation and interest in these subjects. I would also like to thank my undergraduate professors of Biochemistry and Biology Dr. Canfield and Dr. Lopilato, respectively, for their mentorship and sparking my interest in research which propelled me to graduate school.

This thesis could not have been possible without the guidance, mentorship and dedication of my advisor Dr. Pamela Peralta-Yahya. She has helped develop me into a better scientist and has provided career development opportunities for which I am forever grateful. I would also like to thank members of the Peralta-Yahya Lab from my tenure starting with my peer mentors Amy Ehrenworth and Stephen Sarria who paved the way before me and were there with me through the highs and lows of research. I am glad to still call them my friends. I would also like to thank Post-Docs Kuntal Mukherjee and Rebecca

Howie for always providing good advice in lab. Next, I would like to thank Nick Kruyer for always having a great attitude and being a great person to bounce ideas off of.

The work encompassed here could not have been completed without the help of wonderful research technicians in the Peralta-Yahya Lab. Sara Eisen has been my right hand in keeping projects organized and on task and has become a great friend. Widianti Sugianto has been with the Peralta-Yahya Lab since she was an undergrad and I have enjoyed watching her evolve into a brilliant scientist. Aurelia Allen, Anita Minniefield, Kaitlyn Hoover, Paul Branham, and Hanfei Wang were all vital in producing this research and I am grateful for their help and friendship. I would also like to thank Elise Gowen for being a good sport and also willing to help me out when I needed it during the last few months. Lastly, I would like to thank the undergraduates who I have worked with in the Peralta-Yahya Lab, especially Lily Torp. Lily brightens everyone's day and it is a joy to share the trials and tribulations of research with her. I have learned something from working with each person in the Peralta-Yahya Lab and thank everyone for making it fun to go into work every day.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

3PB	3-phenylbutyr-aldehyde
5-HT	serotonin
5-HTR$_4$	serotonin receptor 4
β2AR	β2-adrenergic receptor
A$_{2A}$R	adenosine receptor
CCR	chemokine receptor
DAG	diacylglycerol
CS	chemical space
D$_4$	dopamine 4 receptor
DMEM	Dulbecco's modified eagle medium
DMSO	dimethyl sulfoxide
exOR	ectopically expressed olfactory receptor
FDA	U.S. Food and Drug Administration
GECI	genetically encoded calcium indicators
GFP	green fluorescent protein
GI	gastrointestinal
GPCR	G protein-coupled receptor
HTS	high-throughput screening
IBS	irritable bowel syndrome
IBS-C	irritable bowel syndrome with constipation
IP$_3$	inositol triphosphate
MAPK	mitogen-activated protein kinase

MS	mass spectrometry
OR	olfactory receptor
PC	principal component
PRESTO	parallel receptorome expression and screening via transcriptional output
SMILES	simplified molecular-input line-entry system
SPR	surface plasmon resonance
SSTR2	human somatostatin receptor 2
TBB	*tert*-butylbenzene
TEV	tobacco etch virus

SUMMARY

G protein-coupled receptors (GPCRs) are the largest family of proteins expressed in humans and are the target of over one-third of U.S. Food and Drug Administration (FDA) approved drugs. Surprisingly, many GPCRs are "orphans" or lack endogenous ligands that prevent further study. Traditionally, GPCR-based assays used for drug discovery and deorphanization have relied on low-throughput methods and non-optimal assay hosts. Here, we present recent advances in GPCR-based assays and expand GPCR-based assays in the yeast, *Saccharomyces cerevisiae*.

In **Chapter 1**, we review the current state of the art of high-throughput GPCR-based assays from the past two years. High-throughput screening (HTS) of GPCRs is used not only in pharmaceutical company drug discovery platforms, but is also regularly used to deorphanize, screen or repurpose other drugs, or look for off-target or potential side effects of therapeutics. There are multiple types of assays, such as: β-arrestin recruitment-based, secondary messenger-based (i.e., cAMP, Ca^{2+}, IP_3, etc.), yeast-based, *in vitro*-based, and *in silico*-based assays. Multiple assays exist because there is no one perfect solution for the complexity seen in expression, G protein coupling, signaling, and regulation mechanisms in GPCRs. We highlight key advances in assay development and provide a future outlook on GPCR HTS.

In **Chapter 2**, we develop an olfactory receptor- (OR-)based screening workflow in yeast, screen seven ORs against 57 chemicals and deorphanize two ORs for the first time. There are approximately 400 ORs expressed in humans and over 300 are still orphans. ORs are found ectopically throughout the body and have clinical impact. Of the nine highly

expressed ORs of the human colon, only two are well studied and known endogenous ligands. To further understand ectopic ORs in the colon, we aimed to deorphanize the seven understudied receptors by rapidly creating sensors that coupled OR activation to green fluorescent protein (GFP) expression via the mating pathway in yeast and screening against odorants. Fluorescence microscopy revealed ORs were expressed in yeast and are trafficked to the membrane. Screening led to a total of six OR-chemical pairs that were validated by dose-response curves. This work will advance downstream effects of ectopic colon ORs by providing ligands to study receptor signaling. Additionally, this workflow can be applied to other GPCRs and chemical libraries to quickly identify ligands.

In **Chapter 3**, we optimize the yeast-based sensor for the detection of serotonin with the serotonin receptor 4 (5-HTR_4) and screen 1200 chemicals to find novel ligands. 5-HTR_4 is mainly expressed throughout the colon, where over 90% of serotonin is found, and is implicated in irritable bowel syndrome (IBS). We initially hypothesized that microbial metabolites could be interacting with 5-HTR_4. To test this theory, we screened a natural products library against a yeast-based 5-HTR_4 sensor. We engineered the sensor to improve speed and fold change by replacing GFP with NanoLuc luciferase. Our initial hit was an anti-infective, which propelled the next screen of an anti-infective library. In total, we found three antimicrobial compounds that activated 5-HTR_4 in our yeast-based assay and also promoted wound healing or proliferation in colon epithelium cells that endogenously express 5-HTR_4. This work is intriguing as it may show that antimicrobials not only interact with microbes, but may influence the human host.

Finally, **Chapter 4** contains conclusions and future outlooks for yeast-based assays and **Appendix A** is a detailed method of how to perform the yeast-based serotonin receptor 4 luciferase assay.

CHAPTER 1. ADVANCES IN HIGH-THROUGHPUT G PROTEIN-COUPLED RECEPTOR ASSAYS

1.1 Abstract

G protein-coupled receptors (GPCRs) are the largest family of human proteins and the target of over 30% of FDA approved drugs. High-throughput screening (HTS) of clinically relevant GPCRs is a first-line drug discovery effort in biomedical research. In this opinion, we review recent advances in GPCR HTS. Although our major focus is cell-based assays, we also highlight recent advances in *in vitro* assays using purified receptors, and computational approaches for GPCR HTS. We conclude that as automation equipment becomes more widely available, GPCR HTS will begin to be used to probe basic biological process that will have an outsized impact on personalized medicine. Virtual docking now allows the screening of thousands of molecules in minutes, expediting the transition from drug candidate to pharmaceutical. However, reliance on the crystal structure of the target GPCRs limits the broad applicability of this strategy as only 65 of the 800 druggable GPCRs have structures.

1.2 Introduction

G protein-coupled receptors (GPCRs) are the largest family of human proteins and the target of over 30% of FDA approved drugs [1]. High-throughput screening (HTS) of clinically relevant GPCRs is a first-line drug discovery effort in biomedical research [2]. For example, Pfizer identified maraviroc as an agonist of chemokine receptor 5 (CCR5) and is now an antiviral HIV treatment [3]. GPCR HTS can also be used to identify new

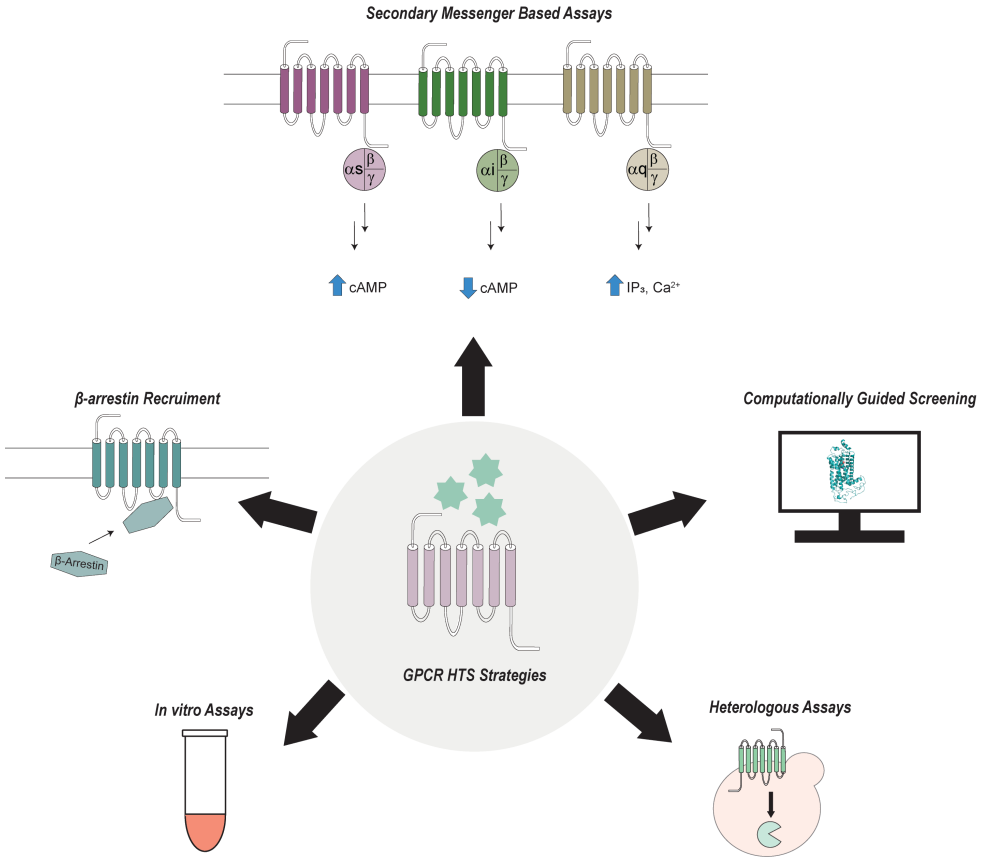

Figure 1.1 Overview of HTS assays for GPCRs.

uses for existing drugs. Raloxifene, an osteoporosis treatment, was identified to also be an inverse agonist of cannabinoid receptor 2 [4], an inflammation regulator. Prazosin, a hypertension treatment, was recognized to also be a partial agonist of CCR3 & CCR4, which are involved in autoimmune diseases and cancer metastasis [5]. About 100 non-olfactory GPCRs remain orphan, i.e., there are no known ligands that activate them, thus their signaling pathways and ultimate phenotypes are unknown [6]. GPCR HTS was used to deorphanize GPR139 by screening ~200,000 chemicals to uncover LP-360924 as an

agonist [7]. Deorphanization of GPR139 has enabled understanding of its role in protecting dopaminergic neurons from degeneration [8], and is a drug target for Parkinson's disease.

The serendipitous expression of human GPCRs outside their endogenous tissue, together with the inefficient transduction of GPCR signaling in heterologous systems hinders the development of GPCR HTS. Briefly, GPCR signaling consists of five steps: 1) a "first messenger" binds a GPCR on the cell surface, 2) the activated GPCR couples to one of 18 different G_α proteins [9] or 4 β-arrestin [10], 3) G_α or β-arrestin amplify the signal by changing the concentration of a "second messenger" or by initiating a phosphorylation cascade, 4) the secondary messenger or terminal phosphorylated protein drives metabolic changes, and 5) the signal ends due to first messenger disappearance, GPCR desensitization, G_α deactivation, second messenger degradation, or terminal protein dephosphoarylation (Figure 1.1).

Here, we review recent advances in GPCR HTS. We classify GPCR assays based on the protein they coupled to and the different secondary messengers they activate. Although we focus on cell-based assays, we highlight advances in *in vitro* assays using purified receptors, and computational approaches for GPCR HTS. Taken together, progress in GPCR HTS continues to reduce the time to arrive at a quality drug candidates. Novel computational approaches now allow the screening of thousands of molecules *in silico*, reducing the number of candidates that need to be tested experimentally.

1.3 Cell-based GPCR high-throughput assays

1.3.1 β-arrestin Recruitment-based Assays

Nearly all GPCRs will recruit *β*-arrestin to initiate receptor desensitization via internalization and some GPCRs recruit *β*-arrestin to initiate a mitogen-activated protein kinase (MAPK) pathway. In the Tango assay, a GPCR fused to a transcription factor via a TEV protease cleavage site recruits a *β*-arrestin linked to a TEV protease upon activation. This recruitment results in the release of the transcription factor and initiation of reporter gene transcription [11]. In 2015, Tango was expanded to more than 300 GPCRs by appending the C-terminus of the vasopressin receptor ("V2 tail") to the C-terminus of GPCRs to generalize the recruitment of *β*-arrestin in PRESTO-Tango [12]. This system was used to screen 91 orphan GPCRs against 446 chemicals to successfully deorphanize two GPCRs: MRGPRX4, with the K_{ATP}-channel blocker nateglindide, and the BB3 receptor, with the antiviral saquinavir [12]. In its most recent application, PRESTO-Tango was used to screen 320 non-olfactory receptors in parallel to identify selective agonists for the D_2 dopamine receptor, a prominent neuropsychiatric drug target [13]. Although Tango has been primarily used as a high-throughput screen to identify GPCR agonists, antagonists can also be identified by running the assay in the presence of an agonist and measure the reduction in signal upon addition of potential antagonists. For example, the *κ*-opioid receptor antagonist ML350 was identified via HTS of over 16,000 chemicals by displacing the agonist U-50488 [14]. *K*-opioid receptor antagonists are used for the treatment of depression.

In 2019, the screening reliability of Tango was improved in PiggyBac-Tango by integrating the GFP reporter and *β*-arresting constructs into the genome [15]. Using PiggyBac-Tango, the dopamine receptor 2, a target for psychiatric disorders and

Parkinson's disease, was screened against 1,000 natural products and wilfortrine was identified as an agonist for the first time [15].

Tango assays link one GPCR to one transcription factor to activate a single reporter gene. To activate multiple genes, ChaCha replaces the transcription factor with a nuclease-deactivated Cas9 and introduces multiple guide RNAs to activate multiple targets simultaneously [16]. ChaCha was used to activate the hM3D receptor to express both interleukin 2 (IL2) and interferon gamma (IFNG) and elicit an immune response. ChaCha is a great platform for comprehensively controlling metabolic fate, however, it has not yet been used for GPCR HTS applications.

Reporters used in β-arrestin recruitment-based assays include luciferase, green fluorescent protein (GFP), split luciferase and beta-lactamase. Although luciferase allows the use of luminescence plate readers, it requires cell lysis of mammalian cells to achieve high signal-to-background ratio. GFP has a long maturation time and requires a high-throughput cell imaging system for signal acquisition in mammalian cells. Split luciferase enables continuous monitoring of GPCR activation [17], but has yet to be used in GPCR HTS applications. In 2018, RNA barcodes read by RNA-seq were introduced as reporters [18]. Specifically, RNA barcodes allow cells expressing different GPCRs to be pooled and analyzed in parallel. In proof-of-principle work, nineteen aminergic receptors across five classes were assigned specific barcodes, pooled and tested against a mix of their five endogenous ligands. All receptors performed equally well in a pooled receptor and chemical environment. RNA barcodes should be used for GPCR HTS of large compound libraries for the identification of novel drugs in the future.

1.3.2 Secondary messenger-based assays

GPCRs couple to four families of G_α subunits ($G_{\alpha s}$, $G_{\alpha i}$, $G_{\alpha q}$ and $G_{\alpha 12/13}$) with 18 different isoforms that result in the upregulation of only a handful of secondary messengers. $G_{\alpha s}$ upregulates cAMP production, $G_{\alpha i}$, inhibits cAMP production, $G_{\alpha q}$, activates phospholipase C, which in turn produce diacylglyercol (DAG) and inositol triphosphate (IP_3), the latter of which binds calcium channels, resulting in Ca^{2+} accumulation, and $G_{\alpha 12/13}$ increases RhoA activity. Here, we will highlight the most widely studied G_α subunits, $G_{\alpha s}$, $G_{\alpha i}$, $G_{\alpha q}$, as direct measurement of $G_{\alpha 12/13}$ signaling is limited [19] due to cross-talk concerns because $G_{\alpha q}$ and $G_{\beta\gamma}$ can also activate RhoA.

1.3.3 cAMP-based assays

Detection of cAMP levels occurs via binding to the CREB transcription factor, which activates transcription of a reporter gene under control of the CRE promoter (pCRE). In 2019, the cAMP-based assay was used to screen β2-adrenergic receptor (β2AR) against ~7,000 chemicals in neuroepithelial cells confirming that β2AR is activated by higenamine, a treatment for heart failure [20]. It is worth noting that GPCRs couple to different proteins depending on the ligand they bind, a concept termed biased signaling [21]. β2AR activates $G_{\alpha s}$, $G_{\alpha i}$ or β-arrestin [22], thus a β2AR ligand identified via cAMP accumulation preferentially activates signaling via $G_{\alpha s}$. To identify ligands that preferentially activate β2AR via β-arrestin or via $G_{\alpha i}$, a Tango assay or a system that reports on the decrease in the cAMP concentration would be needed, respectively. RNA barcodes have also been used to increase the throughput of cAMP-based assays. Specifically, approximately 39

murine olfactory GPCRs were screened against 181 odorants in a pooled assay and 15 ORs were deorphanized [23].

In 2020, the $G_{\alpha i}$ coupled apelin receptor, a drug target for heart failure, was screened against a Bristol Myers Squibb proprietary library and was activated by aryl hydroxy pyrimidinone (AHPs) agonists [24]. AHPs are a new class of apelin agonists identified by HTS that can be used to treat heart failure. When testing a $G_{\alpha i}$ coupled GPCR, addition of an adenylate cyclase activator, such as forskolin, should be used to increase cAMP levels before adding a potential antagonist to measure reduction in cAMP or cAMP-dependent reporter gene transcription [25].

1.3.4 Calcium-based assays

Activation of $G_{\alpha q}$ ultimately results in an increase in Ca^{2+} levels. Calcium dyes, such as Fluo-4, were the original choice to report on calcium fluctuation. However, their cost ($25/96-well plate) and additional assay time (~1 hour) limited their use. In the early 2000s, genetically encoded calcium indicators (GECI) achieved similar affinity as synthetic dyes (~300 nM [26]) and improved fluorescence output [27] that made intracellular calcium detection widespread.

In 2018, the genetically encoded Ca^{2+} and cAMP biosensors were introduced for GPCR HTS [28]. Briefly, a Nomad biosensor is attached to the plasma membrane and consists of a fluorescent protein link to a peptide that undergoes a conformational change upon chemical binding. Specifically, calmodulin-binding peptide (CaM) is used to bind Ca^{2+} and human thyroid A-kinase anchoring protein (HT31) is used to bind cAMP. Upon GPCR activation Ca^{2+} and cAMP can be detected by CaM and HT31, respectively, undergo a

conformational change resulting in increased fluorescence and vesicularization of the sensor, providing a phenotypic output seen as small green dots under the microscope when the cells are activated. Due to its mode of action, Nomad biosensors can detect multiple second messengers simultaneous, thus enabling the screening of chemicals for biased GPCR signaling. A multiplexed Nomad biosensor was used to screen the endothelin B receptor, a $G_{\alpha s}$ and $G_{\alpha q}$ coupled GPCR, against 1200 chemicals. This work elucidated three new ligands with the Ca^{2+} Nomad biosensor ($G_{\alpha q}$ signaling), and one of those ligands also was shown to activate cAMP Nomad ($G_{\alpha s}$ signaling). Taken together, the multiplexed Nomad biosensor has the potential to optimize the screening time of GPCRs with suspected biased signaling. Cytotoxicity of GECIs [29] results in reporter loss and reduced GECIs utility. By physically linking GECI to an antibiotic resistance marker using a self-cleaving peptide (GCaMP6s-2A-Bsr), loss of the GECI results in cell death, thus improving the GECI's fluorescence intensity and long term expression stability [30]. GCaMP6s-2A-Bsr performed on par with Fluo-4 in a 384-well plate format using with the $G_{\alpha q}$ coupled muscarinic acetylcholine as a proof-of-principle GPCR.

Before an increase in Ca^{2+} levels, $G_{\alpha q}$ activation results in an increase in DAG and IP_3 levels, which rapidly metabolizes into inositol diphosphate (IP_2) and inositol monophosphate (IP_1). In 2018, a commercial fluorescence resonance energy transfer (FRET)-based IP_1 kit (IP-One) was used to screen STC-1 cells against casein hydrolysates and found LFC25 increases glucagon-like peptide 1 (GLP-1) production [31]. While not directly measuring one GPCR signal transduction, GLP-1 secretion is modulated by multiple $G_{\alpha q}$ GPCRs. Several DAG fluorescent protein-based sensors are available, but high resolution microscopy is needed, thus limiting its throughput [32]. In 2016, a DAG

FP-sensor was optimized by testing brighter GFP variants for potential use in HTS [33] and is commercially available (DAG Assay Kit). To our knowledge, the DAG FP-based sensor has not been used in HTS of a $G_{\alpha q}$ coupled GPCR.

1.4 *Saccharomyces cerevisiae* as a heterologous host

Mammalian cells express tens of GPCRs on the cell surface (e.g., 75 in HEK293 [34]), changing the concentration of only a handful of second messengers. Thus, changes in reporter gene expression may not be necessarily due to activation of the desired receptor, but activation of a different receptor that results in the measured output. The yeast *Saccharomyces cerevisiae* with only two GPCRs on its cell surface is the ideal host to screen human GPCR activation. Briefly, *S. cerevisiae* expresses only the pheromone (Ste2/Ste3) and a glucose sensing (GPR1) GPCR [35], has no β-arrestins, and GPCR transduction occurs via two G_{α}s, GPA1 and GPA2, which activate either a MAPK pathway or cAMP cascade, respectively [36].

In 2017, the first β-arrestin-based assay was performed in *S. cerevisiae* by fusing the somatostatin receptor 2 (SSTR2) and human β-arrestin to complementary fragments of luciferase [37]. The assay was validated using two known agonist, assay signal was optimal after 60 minutes, albeit it only had a three-fold increase in signal after activation. A drawback of β-arrestin-based assays in yeast is the lack of GPCR kinases often needed to phosphorylate the GPCR for β-arrestin recruitment [38]. Although SSTR2 did not need phosphorylation [37], to expand β-arrestin-based assays in yeast, GPCR kinases will need to be co-expressed.

Human GPCRs can couple to GPA1 or GPA1/human G_α chimeras [39] to hijack and activate the mating pathway to report on GPCR activation. To do this, deletion of the pheromone receptor (*ste2*), along with *far1* and *sst2* that prevents cell cycle arrest and G_α desensitization, respectively, is needed. Leftkowitz et al. [40] originally introduced this strategy in 1990 by coupling β2AR to the yeast MAPK pathway. Of note, yeast GPCR HTS allows use of growth as the reporter, opening the door to screening 10^6-10^8 compounds against one GPCR or 10^6-10^8 GPCR mutants against one chemical every three days. This throughput has been used to engineer GPCRs exclusively activated by inert drugs, a widely used chemogenetic tool in neuroscience [41].

Yeast-based GPCR HTS have been used for drug discovery. In 2015, GPR68, a proton-sensing GPCR involved in asthma, was screened against 446 chemicals using a liquid growth assay read every three days in 96-well plates [42]. Lorazepam, an anti-anxiety medication, was identified as a GPR68 allosteric agonist as it was modulated by pH. Lorazepam was then used to probe GPR68 downstream signaling pathways. To increase the throughput, the luciferase receptor was introduced for GPCR HTS for drug discovery. In 2019, the serotonin 4 receptor (5-HTR_4) was screened against 1,200 natural products and anti-infection compounds where antimicrobials were identified as 5-HTR_4 agonists for the first time [43]. The total assay time was 2.5 hours, which can be performed in parallel, and using a 96-well plate luminescence plate reader the chemical evaluation throughput was one compound per second. The assay used GPA1 to couple to the mating pathway, which is a good mimic for human $G_{\alpha s}$, which 5-HTR_4 couples to, however, GPA1 may not be a good substitute for other $G_{\alpha s}$ [44].

In addition to non-sensory GPCRs, olfactory receptors (ORs) have been coupled to the yeast mating pathway [45]. In 2019, seven ORs were screened against 57 chemicals resulting in the deorphanization of two ORs [46]. The assay linked OR activation to GFP expression, which takes four hours to achieve robust GFP signal. The assay is read using a HT-flow cytometer at a throughput of one 96-well plate per hour.

1.5 *In vitro* GPCR high-throughput assays

The low stability of purified GPCRs makes their screening outside cells or membrane systems challenging. Advances in engineering GPCRs for protein stability [47] have increased the number of *in vitro* GPCR assays. Radioligand binding assays are still used as a sensitive assay to determine GPCR cell surface expression/internalization [48] and ligand binding [49]. However, fluorescence-based techniques along with affinity-based screens are becoming more popular due to the lack of radioactivity and increased throughput.

Fluorescence polarization can distinguish an unbound fluorescently labelled ligand from a receptor-ligand pair by the difference in the molecular tumbling speed. A key challenge with this technique is the availability of fluorescently-labelled substrates. In 2019, neurotensin receptor type 1 (NTS1) with improved stability was used to develop a high-throughput competitive binding assay using a commercially available fluorescently-labeled NTS1 peptide ligand [50]. In a 384-well plate format, the displacement of fluorescently-labeled NTS1 peptide ligand was used to screen over 1,200 chemicals in just one hour of incubation. Surface plasmon resonance (SPR) confirmed 8 of the 15 identified hits.

Affinity MS is an enrichment method that can isolate and identify GPCR ligands from a pool [51]. In 2019, an adenosine receptor ($A_{2A}R$) with improved stability [52] was purified and screened against 20,000 chemicals. First, $A_{2A}R$ was mixed with the chemical library and the unbound chemicals were washed. Next, the receptor was denatured, the chemicals recovered, and incubated with freshly prepared $A_{2A}R$. Nine new $A_{2A}R$ ligands were validated using SPR. Affinity MS can also be applied to membrane systems, albeit they have a higher signal to background ratio than purified enzymes. While affinity mass spectrometry is a sensitive technique, however the requirement of purified enzyme limits it to highly stable GPCRs.

1.6 *In silico* GPCR-based high-throughput assays

With a seemingly endless chemical space to identify GPCR ligands, computational and bioinformatics methods effectively reduce the number of ligand candidates that need to be evaluated experimentally.

Bioinformatic techniques, including sequence alignment, evolutionary trace analysis, and principal component analysis, were used to guide a deorphanization attempt of putative peptide orphan GPCRs leading to the pairing of five receptors with multiple endogenous peptide ligands in 2019 [53]. Key features in peptide GPCRs were identified to narrow the GPCR space to down 21 putative peptide receptors. To reduce the ligand space, secreted proteins were mined from the human proteome to reach 1,227 potential peptide candidates and using machine learning they were focused to 131 peptide candidates. As the downstream signaling pathway for orphan receptors are not known, to screen the 131 peptides against 21 receptors, three nonspecific G-protein signaling assays were used: mass

redistribution, receptor internalization (time-resolved-FRET), and β-arrestin recruitment (Tango). Given that computational approaches significantly reduce the number of GPCR-ligand hits to be validated, real-time lower-throughput approaches, such as mass distribution and receptor internalization, can be used.

Computational docking methods (e.g., Dock3.7 [54]) using ultra-large chemical libraries expand chemical library diversity and size from thousands, e.g., human proteome ~20,000 to hundreds of millions, e.g., potential small molecule drug candidates 10^{63} [55]. In 2019, 138 million chemicals were virtually screened against the dopamine 4 receptor (D_4), the pharmacological target for schizophrenia and Parkinson's disease, in a single day [56]. The top 549 computationally generated chemical hits were chemically synthesized, and a radioactivity assay was used to validate 22% of the hits. Additionally, 81 chemicals showed K_is in low the nM to μM range indicating virtual docking was successful in narrowing down millions to a couple of hundred leads to validate experimentally. In 2020, over 150 million chemicals were virtually screened against melatonin receptor 1 (MT1) [57]. After manual inspection of top scoring results, 40 chemicals were synthesized 15 of which showed activity for either melatonin receptor 1 or 2 (61% similar) using Tango and cAMP assays. Of note, the authors identified four new specific agonists for MT1 without the need to search analogous structures further. Online chemical databases that are compatible with virtual screening (e.g., ZINC [58], ChEMBL [59]) have expanded diversity by increasing the sheer scale of chemical testing and by allowing the frequent incorporation of new scaffolds [56]. Use of these ultra-large chemical libraries for virtual docking are most effective when crystal structures are available [60]. Homology modeling and sequence alignment are not useful when only 64 unique GPCR structures are known

[61]. This is a setback in virtual docking, as 85% of non-olfactory GPCRs, and all olfactory GPCRs, do not have structures.

1.7 Future Outlook

To identify new ligands for GPCRs, there is not a single go-to assay. GPCRs couple through four different G_α subunits in addition to β-arrestin, thus the GPCR HTS chosen needs to take into account the identity of the GPCR as well as the protein it couples for signal transduction. GPCR HTS are robust and have led to the identification of new drugs, the repurposing of existing drugs, and more recently the identification of chemicals with a G-protein transduction bias. With the wider availability of automation equipment together with more user-friendly computational platforms, we are at an apogee of GPCR HTS.

To date, GPCR HTS applications have focused on drug discovery. However, as automation equipment becomes more common in academic laboratories, we foresee GPCR HTS being used to question and probe basic biological processes. For example, different GPCR isoforms tend to be expressed in different cell types. By applying GPCR HTS, we can start to elucidate the extent different cell types are activated by different ligands. Further, different patient populations tend to have different GPCR alleles, and GPCR HTS could help elucidate the extent to which different GPCR alleles are activated. With the potential of having each patient's genome sequenced in the near future, we can envision GPCR HTS of chemical and peptide libraries to be a common occurrence to generate lifesaving therapies.

Bioinformatics and virtual docking presents another unique way for the field to rapidly deorphanize GPCRs or find potential drugs. However, there is a higher success rate when

the docking analysis is done in the specific GPCR structure and not a homolog thereof. With 64 unique GPCR structures and a total of 800 druggable GPCRs, we have some way to go to apply virtual docking strategies to all GPCRs. To date, machine learning has been used to reduce the chemicals search space. In the future, machine learning promises to provide an avenue to combine bioinformatics, experimental and docking data to predict potential GPCR ligands and even deorphanize GPCRs in the future.

1.8 References

[1] Hauser, A. S., Attwood, M. M., Rask-Andersen, M., Schioth, H. B., and Gloriam, D. E. (2017) Trends in GPCR drug discovery: new agents, targets and indications, *Nat. Rev. Drug Discov. 16*, 829-842.

[2] Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U., and Sittampalam, G. S. (2011) Impact of high-throughput screening in biomedical research, *Nat. Rev. Drug Discov. 10*, 188-195.

[3] Dorr, P., Westby, M., Dobbs, S., Griffin, P., Irvine, B., Macartney, M., Mori, J., Rickett, G., Smith-Burchnell, C., Napier, C., Webster, R., Armour, D., Price, D., Stammen, B., Wood, A., and Perros, M. (2005) Maraviroc (UK-427,857), a potent, orally bioavailable, and selective small-molecule inhibitor of chemokine receptor CCR5 with broad-spectrum anti-human immunodeficiency virus type 1 activity, *Antimicrob. Agents Ch 49*, 4721-4732.

[4] Kumar, P., and Song, Z. H. (2013) Identification of raloxifene as a novel CB2 inverse agonist, *Biochem. Biophys. Res. Commun. 435*, 76-81.

[5] Gao, X. L., Abdelkarim, H., Albee, L. J., Volkman, B. F., Gaponenko, V., and Majetschak, M. (2018) Partial agonist activity of alpha(1)-adrenergic receptor antagonists for chemokine (C-X-C motif) receptor 4 and atypical chemokine receptor 3, *PloS one 13*.

[6] Alexander, S. P. H., Christopoulos, A., Davenport, A. P., Kelly, E., Mathie, A., Peters, J. A., Veale, E. L., Armstrong, J. F., Faccenda, E., Harding, S. D., Pawson, A. J., Sharman, J. L., Southan, C., Davies, J. A., and Collaborators, C. (2019) THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: G protein-coupled receptors, *Br. J. Pharmacol. 176* Suppl 1, S21-S141.

[7] Hu, L. Y. A., Tang, P. M., Eslahi, N. K., Zhou, T., Barbosa, J., and Liu, Q. Y. (2009) Identification of Surrogate Agonists and Antagonists for Orphan G-Protein-Coupled Receptor GPR139, *J. Biomol. Screen 14*, 789-797.

[8] Andersen, K. B., Johansen, J. L., Hentzer, M., Smith, G. P., and Dietz, G. P. H. (2016) Protection of Primary Dopaminergic Midbrain Neurons by GPR139 Agonists Supports Different Mechanisms of MPP+ and Rotenone Toxicity, *Front. Cell. Neurosci. 10*.

[9] Syrovatkina, V., Alegre, K. O., Dey, R., and Huang, X. Y. (2016) Regulation, Signaling, and Physiological Functions of G-Proteins, *J. Mol. Biol. 428*, 3850-3868.

[10] van Gastel, J., Hendrickx, J. O., Leysen, H., Santos-Otte, P., Luttrell, L. M., Martin, B., and Maudsley, S. (2018) beta-Arrestin Based Receptor Signaling Paradigms:

Potential Therapeutic Targets for Complex Age-Related Disorders, *Front. Pharmacol. 9*, 1369.

[11] Barnea, G., Strapps, W., Herrada, G., Berman, Y., Ong, J., Kloss, B., Axel, R., and Lee, K. J. (2008) The genetic design of signaling cascades to record receptor activation, *Proc. Natl. Acad. Sci. U.S.A. 105*, 64-69.

[12] Kroeze, W. K., Sassano, M. F., Huang, X. P., Lansu, K., McCorvy, J. D., Giguere, P. M., Sciaky, N., and Roth, B. L. (2015) PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome, *Nat. Struct. Mol. Biol. 22*, 362-U328.

[13] Fan, L., Tan, L., Chen, Z., Qi, J., Nie, F., Luo, Z., Cheng, J., and Wang, S. (2020) Haloperidol bound D2 dopamine receptor structure inspired the discovery of subtype selective ligands, *Nat. Commun. 11*, 1074.

[14] Guerrero, M., Urbano, M., Brown, S. J., Cayanan, C., Ferguson, J., Cameron, M., Devi, L. A., Roberts, E., and Rosen, H. (2010) Optimization and characterization of an opioid kappa receptor (OPRK1) antagonist, In Probe Reports from the NIH Molecular Libraries Program, Bethesda (MD).

[15] Li, F., Jiang, X., Luo, L. L., Xu, Y. M., Huang, X. X., Huang, C., and Zhang, Y. (2019) A piggyBac-based TANGO GFP assay for high throughput screening of GPCR ligands in live cells, *Cell Commun. Signal. 17*.

[16] Kipniss, N. H., Dingal, P. C. D. P., Abbott, T. R., Gao, Y. C., Wang, H. F., Dominguez, A. A., Labanieh, L., and Qi, L. S. (2017) Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system, *Nat. Commun. 8*.

[17] Littmann, T., Buschauer, A., and Bernhardt, G. (2019) Split luciferase-based assay for simultaneous analyses of the ligand concentration- and time-dependent recruitment of beta-arrestin2, *Anal. Biochem. 573*, 8-16.

[18] Galinski, S., Wichert, S. P., Rossner, M. J., and Wehr, M. C. (2018) Multiplexed profiling of GPCR activities by combining split TEV assays and EXT-based barcoded readouts, *Sci. Rep. 8*.

[19] Masuho, I., Skamangas, N. K., and Martemyanov, K. A. (2019) Live cell optical assay for precise characterization of receptors coupling to Galpha12, *Basic Clin. Pharmacol. Toxicol.*

[20] Chen, Y., Guo, B., Zhang, H., Hu, L., and Wang, J. (2019) Higenamine, a Dual Agonist for beta 1- and beta 2-Adrenergic Receptors Identified by Screening a Traditional Chinese Medicine Library, *Planta Med. 85*, 738-744.

[21] Smith, J. S., Lefkowitz, R. J., and Rajagopal, S. (2018) Biased signalling: from simple switches to allosteric microprocessors, *Nat. Rev. Drug Discov. 17*, 243-260.

[22] Jean-Charles, P. Y., Kaur, S., and Shenoy, S. K. (2017) G Protein-Coupled Receptor Signaling Through beta-Arrestin-Dependent Mechanisms, *J. Cardiovasc. Pharmacol. 70*, 142-158.

[23] Jones, E. M., Jajoo, R., Cancilla, D., Lubock, N. B., Wang, J., Satyadi, M., Cheung, R., de March, C., Bloom, J. S., Matsunami, H., and Kosuri, S. (2019) A Scalable, Multiplexed Assay for Decoding GPCR-Ligand Interactions with RNA Sequencing, *Cell Syst. 8*, 254-+.

[24] Myers, M. C., Bilder, D. M., Cavallaro, C. L., Chao, H. J., Su, S., Burford, N. T., Nayeem, A., Wang, T., Yan, M., Langish, R. A., Dabros, M., Li, Y. X., Rose, A. V., Behnia, K., Onorato, J. M., Gargalovic, P. S., Wexler, R. R., and Lawrence, R. M. (2020) Discovery and SAR of aryl hydroxy pyrimidinones as potent small molecule agonists of the GPCR APJ, *Bioorg. Med. Chem. Lett. 30*, 126955.

[25] Wang, T., Li, Z., Cvijic, M. E., Zhang, L., and Sum, C. S. (2004) Measurement of cAMP for Galphas- and Galphai Protein-Coupled Receptors (GPCRs), In Assay Guidance Manual (Sittampalam, G. S., Grossman, A., Brimacombe, K., Arkin, M., Auld, D., Austin, C. P., Baell, J., Bejcek, B., Caaveiro, J. M. M., Chung, T. D. Y., Coussens, N. P., Dahlin, J. L., Devanaryan, V., Foley, T. L., Glicksman, M., Hall, M. D., Haas, J. V., Hoare, S. R. J., Inglese, J., Iversen, P. W., Kahl, S. D., Kales, S. C., Kirshner, S., Lal-Nag, M., Li, Z., McGee, J., McManus, O., Riss, T., Saradjian, P., Trask, O. J., Jr., Weidner, J. R., Wildey, M. J., Xia, M., and Xu, X., Eds.), Bethesda (MD).

[26] Paredes, R. M., Etzler, J. C., Watts, L. T., Zheng, W., and Lechleiter, J. D. (2008) Chemical calcium indicators, *Methods 46*, 143-151.

[27] Chen, T. W., Wardill, T. J., Sun, Y., Pulver, S. R., Renninger, S. L., Baohan, A., Schreiter, E. R., Kerr, R. A., Orger, M. B., Jayaraman, V., Looger, L. L., Svoboda, K.,

and Kim, D. S. (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity, *Nature 499*, 295-+.

[28] Mella, R. M., Kortazar, D., Roura-Ferrer, M., Salado, C., Valcarcel, M., Castilla, A., and Villace, P. (2018) Nomad Biosensors: A New Multiplexed Technology for the Screening of GPCR Ligands, *Slas Technol. 23*, 207-216.

[29] Yang, Y., Liu, N., He, Y., Liu, Y., Ge, L., Zou, L., Song, S., Xiong, W., and Liu, X. (2018) Improved calcium sensor GCaMP-X overcomes the calcium channel perturbations induced by the calmodulin in GCaMP, *Nat. Commun. 9*, 1504.

[30] Wu, N., Nishioka, W. K., Derecki, N. C., and Maher, M. P. (2019) High-throughput-compatible assays using a genetically-encoded calcium indicator, *Sci. Rep. 9.*

[31] O'Halloran, F., Bruen, C., McGrath, B., Schellekens, H., Murray, B., Cryan, J. F., Kelly, A. L., McSweeney, P. L. H., and Giblin, L. (2018) A casein hydrolysate increases GLP-1 secretion and reduces food intake, *Food Chem 252*, 303-310.

[32] Sato, M., Ueda, Y., and Umezawa, Y. (2006) Imaging diacylglycerol dynamics at organelle membranes, *Nat. Methods 3*, 797-799.

[33] Tewson, P. H., Martinka, S., Shaner, N. C., Hughes, T. E., and Quinn, A. M. (2016) New DAG and cAMP Sensors Optimized for Live-Cell Assays in Automated Laboratories, *J. Biomol. Screen 21*, 298-305.

[34] Atwood, B. K., Lopez, J., Wager-Miller, J., Mackie, K., and Straiker, A. (2011) Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis, *BMC Genomics 12*, 14.

[35] Versele, M., Lemaire, K., and Thevelein, J. M. (2001) Sex and sugar in yeast: two distinct GPCR systems, *Embo Rep. 2*, 574-579.

[36] Alvaro, C. G., O'Donnell, A. F., Prosser, D. C., Augustine, A. A., Goldman, A., Brodsky, J. L., Cyert, M. S., Wendland, B., and Thorner, J. (2014) Specific alpha-arrestins negatively regulate Saccharomyces cerevisiae pheromone response by down-modulating the G-protein-coupled receptor Ste2, *Mol. Cell. Biol. 34*, 2660-2681.

[37] Fukutani, Y., Ishii, J., Kondo, A., Ozawa, T., Matsunami, H., and Yohda, M. (2017) Split luciferase complementation assay for the analysis of G protein-coupled receptor ligand response in Saccharomyces cerevisiae, *Biotechnol. Bioeng. 114*, 1354-1361.

[38] Reiter, E., and Lefkowitz, R. J. (2006) GRKs and beta-arrestins: roles in receptor silencing, trafficking and signaling, *Trends Endocrinol. Metab. 17*, 159-165.

[39] Price, L. A., Kajkowski, E. M., Hadcock, J. R., Ozenberger, B. A., and Pausch, M. H. (1995) Functional Coupling of a Mammalian Somatostatin Receptor to the Yeast Pheromone Response Pathway, *Mol. Cell. Biol. 15*, 6188-6195.

[40] King, K., Dohlman, H. G., Thorner, J., Caron, M. G., and Lefkowitz, R. J. (1990) Control of Yeast Mating Signal Transduction by a Mammalian Beta-2-Adrenergic Receptor and Gs Alpha-Subunit, *Science 250*, 121-123.

[41] Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S., and Roth, B. L. (2007) Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand, *Proc. Natl. Acad. Sci. U. S. A. 104*, 5163-5168.

[42] Huang, X. P., Karpiak, J., Kroeze, W. K., Zhu, H., Chen, X., Moy, S. S., Saddoris, K. A., Nikolova, V. D., Farrell, M. S., Wang, S., Mangano, T. J., Deshpande, D. A., Jiang, A., Penn, R. B., Jin, J., Koller, B. H., Kenakin, T., Shoichet, B. K., and Roth, B. L. (2015) Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65, *Nature 527*, 477-483.

[43] Yasi, E. A., Allen, A. A., Sugianto, W., and Peralta-Yahya, P. (2019) Identification of Three Antimicrobials Activating Serotonin Receptor 4 in Colon Cells, *ACS Synth. Biol.*

[44] Ehrenworth, A. M., Claiborne, T., and Peralta-Yahya, P. (2017) Medium-Throughput Screen of Microbially Produced Serotonin via a G-Protein-Coupled Receptor-Based Sensor, *Biochemistry 56*, 5471-5475.

[45] Minic, J., Persuy, M. A., Godel, E., Aioun, J., Connerton, I., Salesse, R., and Pajot-Augy, E. (2005) Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening, *FEBS J. 272*, 524-537.

[46] Yasi, E. A., Eisen, S. L., Wang, H., Sugianto, W., Minniefield, A. R., Hoover, K. A., Branham, P. J., and Peralta-Yahya, P. (2019) Rapid Deorphanization of Human Olfactory Receptors in Yeast, *Biochemistry 58*, 2160-2166.

[47] Klenk, C., Ehrenmann, J., Schutz, M., and Pluckthun, A. (2016) A generic selection system for improved expression and thermostability of G protein-coupled receptors by directed evolution, *Sci. Rep. 6*.

[48] Xu, W., Reith, M. E. A., Liu-Chen, L. Y., and Kortagere, S. (2019) Biased signaling agonist of dopamine D3 receptor induces receptor internalization independent of beta-arrestin recruitment, *Pharmacol. Res. 143*, 48-57.

[49] Butkiewicz, M., Rodriguez, A. L., Rainey, S. E., Wieling, J., Luscombe, V. B., Stauffer, S. R., Lindsley, C. W., Conn, P. J., and Meiler, J. (2019) Identification of Novel Allosteric Modulators of Metabotropic Glutamate Receptor Subtype 5 Acting at Site Distinct from 2-Methyl-6-(phenylethynyl)-pyridine Binding, *ACS Chem. Neurosci. 10*, 3427-3436.

[50] Heine, P., Witt, G., Gilardi, A., Gribbon, P., Kummer, L., and Pluckthun, A. (2019) High-Throughput Fluorescence Polarization Assay to Identify Ligands Using Purified G Protein-Coupled Receptor, *Slas Discov. 24*, 915-927.

[51] Lu, Y., Qin, S. S., Zhang, B. J., Dai, A. T., Cai, X. Q., Ma, M. N., Gao, Z. G., Yang, D. H., Stevens, R. C., Jacobson, K. A., Wang, M. W., and Shui, W. Q. (2019) Accelerating the Throughput of Affinity Mass Spectrometry-Based Ligand Screening toward a G Protein-Coupled Receptor, *Anal. Chem. 91*, 8162-8169.

[52] Liu, W., Chun, E., Thompson, A. A., Chubukov, P., Xu, F., Katritch, V., Han, G. W., Roth, C. B., Heitman, L. H., IJzerman, A. P., Cherezov, V., and Stevens, R. C. (2012)

Structural Basis for Allosteric Regulation of GPCRs by Sodium Ions, *Science 337*, 232-236.

[53] Foster, S. R., Hauser, A. S., Vedel, L., Strachan, R. T., Huang, X. P., Gavin, A. C., Shah, S. D., Nayak, A. P., Haugaard-Kedstrom, L. M., Penn, R. B., Roth, B. L., Brauner-Osborne, H., and Gloriam, D. E. (2019) Discovery of Human Signaling Systems: Pairing Peptides to G Protein-Coupled Receptors, *Cell 179*, 895-+.

[54] Coleman, R. G., Carchia, M., Sterling, T., Irwin, J. J., and Shoichet, B. K. (2013) Ligand pose and orientational sampling in molecular docking, *PloS one 8*, e75992.

[55] Bohacek, R. S., McMartin, C., and Guida, W. C. (1996) The art and practice of structure-based drug design: A molecular modeling perspective, *Med. Res. Rev. 16*, 3-50.

[56] Lyu, J., Wang, S., Balius, T. E., Singh, I., Levit, A., Moroz, Y. S., O'Meara, M. J., Che, T., Algaa, E., Tolmachova, K., Tolmachev, A. A., Shoichet, B. K., Roth, B. L., and Irwin, J. J. (2019) Ultra-large library docking for discovering new chemotypes, *Nature 566*, 224-+.

[57] Stein, R. M., Kang, H. J., McCorvy, J. D., Glatfelter, G. C., Jones, A. J., Che, T., Slocum, S., Huang, X. P., Savych, O., Moroz, Y. S., Stauch, B., Johansson, L. C., Cherezov, V., Kenakin, T., Irwin, J. J., Shoichet, B. K., Roth, B. L., and Dubocovich, M. L. (2020) Virtual discovery of melatonin receptor ligands to modulate circadian rhythms, *Nature*.

[58] Irwin, J. J., and Shoichet, B. K. (2005) ZINC--a free database of commercially available compounds for virtual screening, *J. Chem. Inf. Model 45*, 177-182.

[59] Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L. J., Cibrian-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magarinos, M. P., Overington, J. P., Papadatos, G., Smit, I., and Leach, A. R. (2017) The ChEMBL database in 2017, *Nucleic Acids Res. 45*, D945-D954.

[60] Loo, J. S. E., Emtage, A. L., Ng, K. W., Yong, A. S. J., and Doughty, S. W. (2018) Assessing GPCR homology models constructed from templates of various transmembrane sequence identities: Binding mode prediction and docking enrichment, *Journal of molecular graphics & modelling 80*, 38-47.

[61] Pandy-Szekeres, G., Munk, C., Tsonkov, T. M., Mordalski, S., Harpsoe, K., Hauser, A. S., Bojarski, A. J., and Gloriam, D. E. (2018) GPCRdb in 2018: adding GPCR structure models and ligands, *Nucleic Acids Res. 46*, D440-D446.

CHAPTER 2. RAPID DEORPHANIZATION OF HUMAN OLFACTORY RECEPTORS IN YEAST

Reprinted with permission from:

Yasi, E. A., Eisen, S. L., Wang, H., Sugianto, W., Minniefield, A, R., Hoover, K. A., Branham, P. J., and Peralta-Yahya, P. *Biochemistry* **2019,** *58* (16), 2160-2166.

2.1 Abstract

Olfactory receptors are ectopically expressed (exORs) in more than 16 different tissues. Studying the role of exORs is hindered by the lack of known ligands that activate these receptors. Of particular interest are exORs in the colon, the section of the gastrointestinal tract with the greatest diversity of microbiota where ORs may be participating in host−microbiome communication. Here, we leverage a G-protein-coupled receptor (GPCR)-based yeast sensor strain to generate sensors for seven ORs highly expressed in the colon. We screen the seven colon ORs against 57 chemicals likely to bind ORs in olfactory tissue. We successfully deorphanize two colon exORs for the first time, OR2T4 and OR10S1, and find alternative ligands for OR2A7. The same OR deorphanization workflow can be applied to the deorphanization of other ORs and GPCRs in general. Identification of ligands for OR2T4, OR10S1, and OR2A7 will enable the study of these ORs in the colon. Additionally, the colon OR-based sensors will enable the elucidation of endogenous colon metabolites that activate these receptors. Finally, deorphanization of OR2T4 and OR10S1 supports studies of the neuroscience of olfaction.

2.2 Introduction

Olfactory receptors (ORs) make up the largest group of G-protein-coupled receptors (GPCRs) and are used by organisms to sense their chemical environment.[1] Rather than

binding a single chemical, ORs bind a range of chemicals with different affinities.[2] Identification of ligands that activate orphan ORs remains a challenging process.[3] To date, only 10% of the approximately 400 human ORs have known ligands.[3,4] Both experimental and computational approaches have been applied to OR deorphanization. Among the most large-scale ones is the use of a mammalian-based OR assay to screen 394 human ORs against 73 chemicals to deorphanize 18 ORs.[5,6] Computational approaches have been used to generate machine learning algorithms by virtually screening hundreds of chemicals against ORs with known ligands, yet the algorithms have failed to deorphanize ORs.[7]

More than 20% of the human ORs are also expressed ectopically outside the olfactory tissue,[8,9] and their function in these tissues is just starting to be elucidated. OR1D2 expression in the testis has been implicated in chemotaxis;[10] OR51E2 expression in the kidney mediates renin secretion,[11] and OR151E1 and OR51E2 in the colon respond to short-chain fatty acids[11−14] likely produced by gut microbiota leading to changes in gene expression.[13] The key to studying the role of ectopically expressed ORs (exORs) and the identification of endogenous ligands present in the tissues in which they are expressed is the availability of "synthetic ligands" that activate these receptors in the laboratory. To date, nine of the 84 ectopically expressed human ORs have known ligands.[10−18] In the human colon, nine ORs are expressed with high confidence: OR10S1, OR2A7, OR2A42, OR2L13, OR2T4, OR2W3, OR51B5, OR51E1, and OR51E2.[9] The role of OR51E2[14] and OR51E1[19] in the colon has been well studied. Synthetic ligands for OR51B5 (isononyl alcohol[17]), OR2A7 (cyclohexyl salicylate[20]), and OR2W3 (nerol[18]) have been reported.

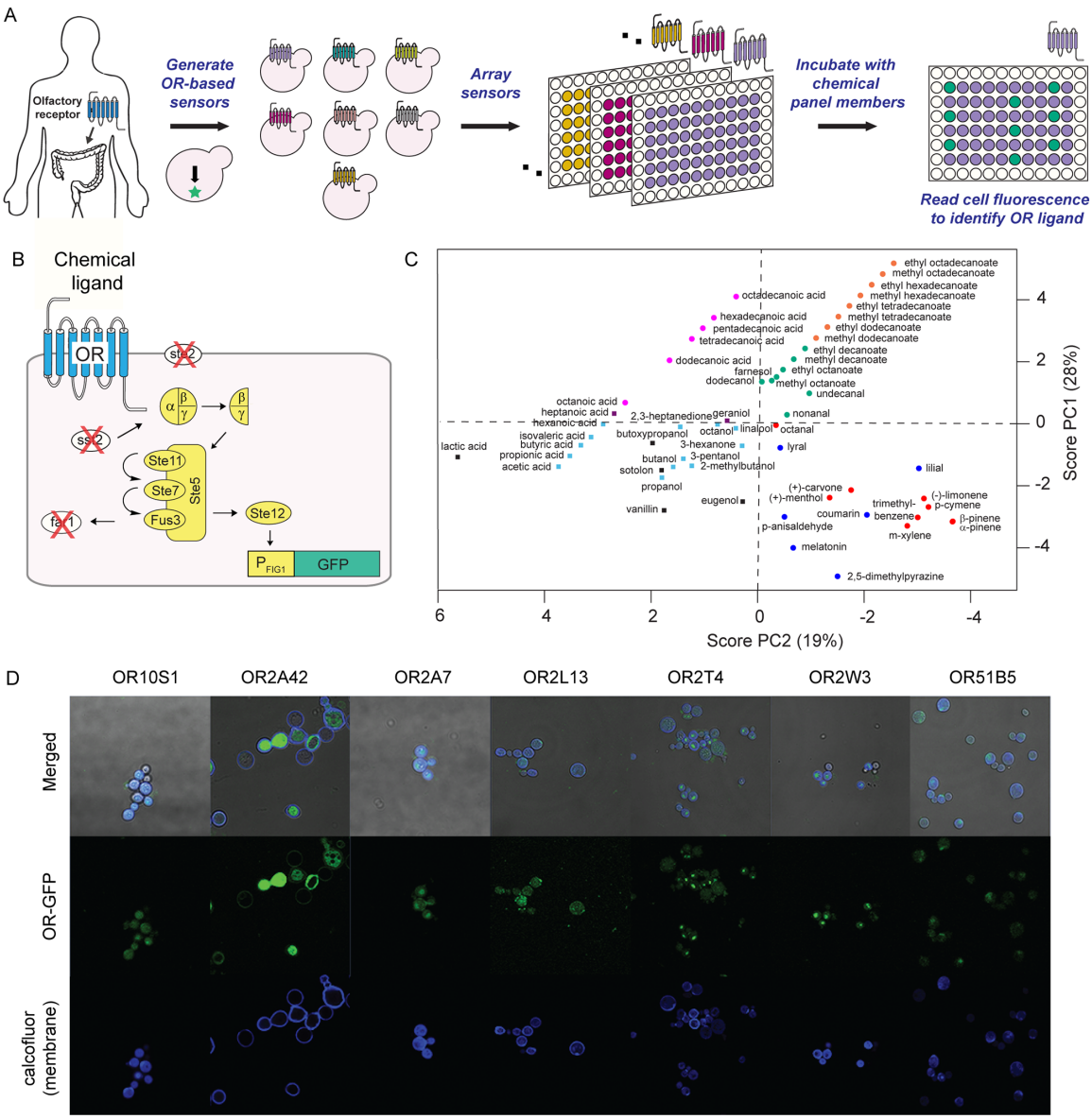

Figure 2.1 – Ectopically expressed olfactory receptor (exOR) sensors and the chemical panel. (A) OR deorphanization workflow. (B) Yeast-based OR sensor. OR (blue) is expressed in a yeast sensor strain that links receptor activation to green fluorescent protein (GFP). (C) Principal component analysis of the 57-member chemical panel. Chemical Spaces (CS): CS1, red dot; CS2, green dot; CS3, blue dot; CS4, orange dot; CS5, purple square; CS6, light blue square; CS7, pink dot; CS8, black square. (D) Localization of human ORs when expressed in yeast.

The other four ORs (OR10S1, OR2A42, OR2L13, and OR2T4) remain orphans. Deorphanization of these four ORs and finding alternative ligands for OR51B5, OR2A7, and OR2W3 would help to elucidate the role of ORs in the colon. Additionally,

deorphanization of OR10S1, OR2A42, OR2L13, and OR2T4 would aid in the studies of the neuroscience of olfaction.

Here, we develop a workflow for the rapid deorphanization of olfactory receptors and apply it to the deorphanization of human ORs expressed in the colon (Figure 2.1). Specifically, we leverage a previously engineered GPCR-based yeast sensor strain[21] (Figure 2.1) to generate seven colon exOR sensors. We screen each of these colon ORs against a 57-member chemical panel to successfully deorphanize two receptors, OR2T4 and OR10S1, and identify two alternative ligands for OR2A7. The OR deorphanization workflow can be readily applied to the deorphanization of other ORs. The yeast-based OR assay is faster than its mammalian cell counterpart due to the shortened doubling time and nonrequirement of cell passage. In addition, yeast-based OR sensors can be stored for up to a month at 4 °C before use. The newly identified ligands for OR2T4, OR10S1, and OR2A7 now allow for the elucidation of the role of these ORs in the colon. The OR high-throughput screening assay enables the elucidation of endogenous OR ligands in the colon.

2.3 Materials and Methods

2.3.1 Materials

Tert-butylbenzene (B90602) and 3-phenylbutyraldehyde (289027) were purchased from Sigma-Aldrich. Nonanoic acid (N0288) and nerol (N0077) were purchased from TCI Chemicals. Supplier information for the chemical panels can be found in Table F.1.

2.3.2 Principal Component Analysis for the Chemical Panel.

Simplified molecular-input line-entry system (SMILES) codes for the 57 chemicals in the panel were obtained from PubChem and used as inputs to Instant JChem (ChemAxon) to acquire information about 23 descriptors. Eight descriptors were chemical functional groups, and 15 descriptors were from a method by Wenderski *et al*.[22] The principal components (PCs) were calculated using Solo Eigenvector (Eigenvector Research). Three PCs, which encompass 63% cumulative variance, were selected by evaluating eigenvalues. PC scores exported from Solo were used to create scatter plots in MATLAB.

2.3.3 Strains and Plasmids.

Human olfactory receptors OR2A4, OR2W3, OR2T4, OR51B5, OR2L13, OR10S1, and OR2A7 were codon-optimized for *Saccharomyces cerevisiae* and commercially synthesized. ORs were cloned into pKM111 at *Bam*HI/*Sac*II sites via Gibson assembly to create pHW3, pHW6, pHW7, pHW18, pPB8, pHW20, and pHW21, respectively. The sequences of the plasmids were verified using primers EY46 and HW12. To construct the OR-based sensors, pHW3, pHW6, pHW7, pHW18, pPB8, pHW20, or pHW21 was co-transformed with pRS415-P_{FIG1}-eGFP-Leu2 (pKM586) into yeast sensor strain PPY140[21] (W303 *Δfar1*, *Δste2*, *Δsst2*) to generate PPY1801−1807. To construct the control strain lacking the OR, PPY140 was co-transformed with pKM586 and an empty vector to generate PPY1800.

To construct OR-GFP fusion plasmids, ORs were amplified from pHW3, pHW6, pHW7, pHW18, pPB8, pHW20, and pHW21 using primer HW1 with HW4, HW7, HW8, HW9, PB89, PB116, and PB118, respectively. GFP was amplified from pKM586 using primer HW12 with overlap to ORs with HW15, HW18, HW19, HW20, PB88, PB115, and PB117. ORs and GFP were cloned into pKM111 at *Bam*HI/*Sac*II sites via Gibson assembly to create pHW22, pHW23, pHW30, pHW33, pPB59, pPB60, and pPB54, respectively.

2.3.4 OR Fluorescence Microscopy.

Overnight cultures of PPY1949−PPY1955 were used to inoculate 20 mL of SD(H^-) until an OD_{600} of 0.06 was reached. After 18 h at 15 °C (150 rpm), cultures were spun down at 3500 rpm for 10 min and resuspended in 200 μL of SD(H^-). One drop (2 μL) of Calcofluor White Stain (Sigma-Aldrich) and one drop (2 μL) of 10% potassium hydroxide were added to the specimen (2 μL) directly on the slide. Yeast was visualized on a Zeiss LSM 700 confocal microscope using the 63× objective lens. GFP was excited using the 488 nm laser line, and Calcofluor white was excited using the 405 nm line.

2.3.5 Screening ORs with a 57-Member Chemical Panel.

Overnight cultures of PPY1801−PPY1807 were used to inoculate 20−40 mL of synthetic complete medium with 2% glucose lacking histidine and leucine [SD(HL^-)] until an OD_{600} of 0.06 was reached. After 18 h at 15 °C (150 rpm), cultures were spun down at 3500 rpm for 10 min and resuspended to an OD_{600} of ≈1 (1/10th of the culture volume). In a 96-well plate, 190 μL of fresh SD(HL^-), 8 μL of the cell suspension, and 2 μL of the solution of the chemical [final chemical concentration of 10 μM, 1% dimethyl sulfoxide (DMSO)] were added. After incubation (4 h, plates covered with Breathe Easy Sealing Membrane, 30 °C, 250 rpm), the GFP fluorescence was read using a Millipore Guava easyCyte HT flow cytometer (λ_{ex} = 488 nm; flow rate of 1.18 μL/s). Samples were run in triplicate. Data from 5000 cells were collected, and 70−95% of viable cells were gated using FlowJo. The geometric mean of the mean fluorescence of the gated cells was used to calculate in Microsoft Excel p values using a Student's t test with two tails using equal variance to define chemical hits ($p < 0.05$).

2.3.6 OR/Chemical Dose-Response Curves.

The same protocol for OR chemical screening was followed. To a test tube were added 4.8 mL of SD(HL^-), 200 μL of the cell suspension, and 50 μL of the chemical (final chemical concentrations of 0−1000 μM, 1% DMSO). After incubation (4 h, plates covered with Breathe Easy Sealing Membrane, 30 °C, 250 rpm), the GFP fluorescence was read using a BD LSR II flow cytometer (488 nm laser line; 515−545 nm filter; FSC, 150 V; SSC, 200 V; FITC, 450 V; FSC threshold, 5000). Samples were run in triplicate. Data from 10000 cells and 70−95% of viable cells were gated using FlowJo. The geometric mean of the mean fluorescence of the gated cells was plotted in OriginPro 2016. Statistically significant points were calculated using Microsoft Excel Student's *t* test with two tails using equal variance. To compare different conditions in a single plot (Figure 2.5 and Figure 2.10), the percent GFP expression was calculated using the formula below. First, the percent GFP expression of every data point for each condition was calculated taking into account GFP (AU) for all conditions to be plotted in the same graph. Then, the percent GFP expression was averaged and the standard deviation calculated.

$$\% \text{ GFP expression} = 100[(GFP_{sample} - GFP_{min})/ (GFP_{max} - GFP_{min})]$$

To calculate EC_{50}s, the data in Figure 2.5 and Figure 2.10 were fitted to a dose-response equation in OriginPro 2016 using the following formula.

$$y = A1 + \frac{A2-A1}{1+10^{(Logx0-x)p}} \qquad EC_{50} = 10^{LOGx0}$$

2.3.7 Chemical toxicity to yeast-based OR sensor.

Overnight cultures of PPY1802, PPY1804, PPY1806, and PPY1807 where diluted to an OD600= 1 in 5 mL of SD(HL-). Cultures were incubated for 4 hours at 30°C, 250 rpm with varying chemical concentrations with a final DMSO concentration of 1%. OD600 was measured after incubation. The experiments were run in triplicate.

2.3.8 mRNA Quantification

Following the OR biosensing protocol, between 2 x 107 and 5 x 107 cells were pelleted and total RNA was extracted using Qiagen RNeasy Mini Kit. RNA concentration was measured using a Nanodrop Lite spectrophotometer and stored at -80°C. Reverse transcription was executed using 700 µg of total RNA using the QuantiTect Reverse Transcription kit (Qiagen). Quantitative PCR reactions were set up using the QuantiTect SYBR Green PCR kit (Qiagen) using cDNA from 42 ng of mRNA and read using a StepOnePlus Real-Time PCR system (Applied Biosciences). Reactions were set up as technical triplicates using primers EY270/EY271 for OR10S1, EY272/EY273 for OR2A7, EY274/EY275 for OR2T4, and ACT-F/ACT-R for actin. OR expression were normalized to the housekeeping gene, ACT1, that encodes actin and relative expression of chemical versus no chemical (DMSO) was compared using the comparative CT method (Livek and Schmittgen, 2001) using the equations below.

$$\Delta C_{T\_Sample} = avgC_{T\_Sample} - avgC_{T\_Actin}$$

$$\Delta\Delta C_T = \Delta C_T - \Delta C_{T\_DMSO}$$

$$2^{-\Delta\Delta C_T} = Mean\ relative\ expression\ of\ gene\ of\ interest\ compared\ to\ DMSO\ control$$

$$2^{-(\Delta\Delta C_T \pm StDev)} = Range\ of\ fold\ change\ in\ expression\ relative\ to\ DMSO\ control$$

$$StDev = standard\ deviation\ of\ \Delta C_{T\_Sample}$$

2.4 Results and Discussion

2.4.1 Structural Diversity in the 57-Member Chemical Panel.

ORs account for 3% of all coding genes in humans[23] and bind a variety of chemicals from terpenes and esters to acids and aldehydes. To identify ligands that activate the seven ORs, we screened them against 57 chemicals likely to bind ORs in olfactory tissue. To understand the structural diversity of the 57 chemicals, each chemical was broken down into 23 descriptors to perform a principal component analysis (Figure 2.1). The top three principal components (PCs) account for 63% of the cumulative variance. On the basis of their PC scores, chemicals can be separated into eight chemical spaces (CSs): CS1, cyclic compounds; CS2, medium-chain esters, alcohols, and aldehydes; CS3, aromatic compounds; CS4, long-chain esters; CS5, heptatonic acid and geraniol; CS6, short-chain hydrocarbons with oxygen-containing functional groups; CS7, long-chain acids; CS8, chemicals with at least two oxygen-containing functional groups. The most well represented chemical space is CS6, accounting for 23% of the chemicals in the panel.

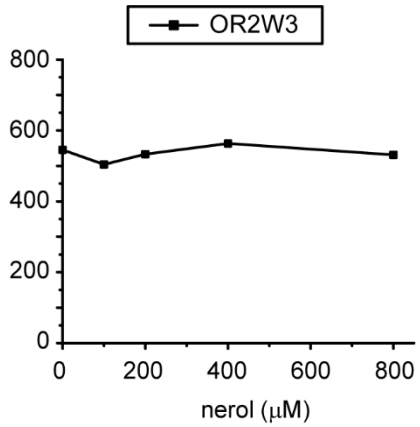

Figure 2.2 Dose-response curve of OR2W3 with nerol.

2.4.2 Colon OR Sensor Generation.

We codon-optimized OR10S1, OR2A7, OR2L13, OR2T4, OR51B5, OR2A42, and OR2W3 for expression in the yeast *S. cerevisiae* and commercially synthesized them. We cloned the ORs under a strong promoter in a high-copy number plasmid and transformed them in the GPCR yeast sensor strain[21] to generate the seven colon OR-based sensors. Incubation of OR2W3 with nerol did not result in a significant increase in the signal after activation (Figure 2.2). ORs sometimes bind different ligands depending on the G_{α} subunit

Table 2.1 – Sequence identity of ectopically expressed olfactory receptors in the colon

	OR2A7	OR2T4	OR2A42	OR2W3	OR51B5	OR2L13	OR10S1
OR2A7							
OR2T4	39.90%						
OR2A42	71.30%	44.00%					
OR2W3	42.70%	41.60%	46.90%				
OR51B5	31.80%	28.80%	32.20%	27.60%			
OR2L13	39.50%	49.20%	42.20%	39.30%	26.20%		
OR10S1	39.60%	36.40%	41.70%	41.60%	28.20%	40.30%	

to which they are coupling.[24,25] Sometimes, an OR is activated by the same ligand independent of the G_{α} subunit to which it is coupled, and when it is coupled to G_{olf}, a signal enhancement in observed.[26,27] Nerol was discovered as a ligand for OR2W3 using G_{olf}.[18] In the yeast system, the ORs couple to the native yeast G_{α} subunit GPA1. It is possible that nerol does not efficiently activate OR2W3 when coupling to GPA1. Cyclohexyl salicylate ($400/mg) and isononyl alcohol ($600/mg) were not readily available and too expensive to be used as synthetic ligands for OR2A7 and OR51B5, respectively. Thus, we set out to identify ligands for all seven ORs. The seven ORs do not have a high degree of sequence

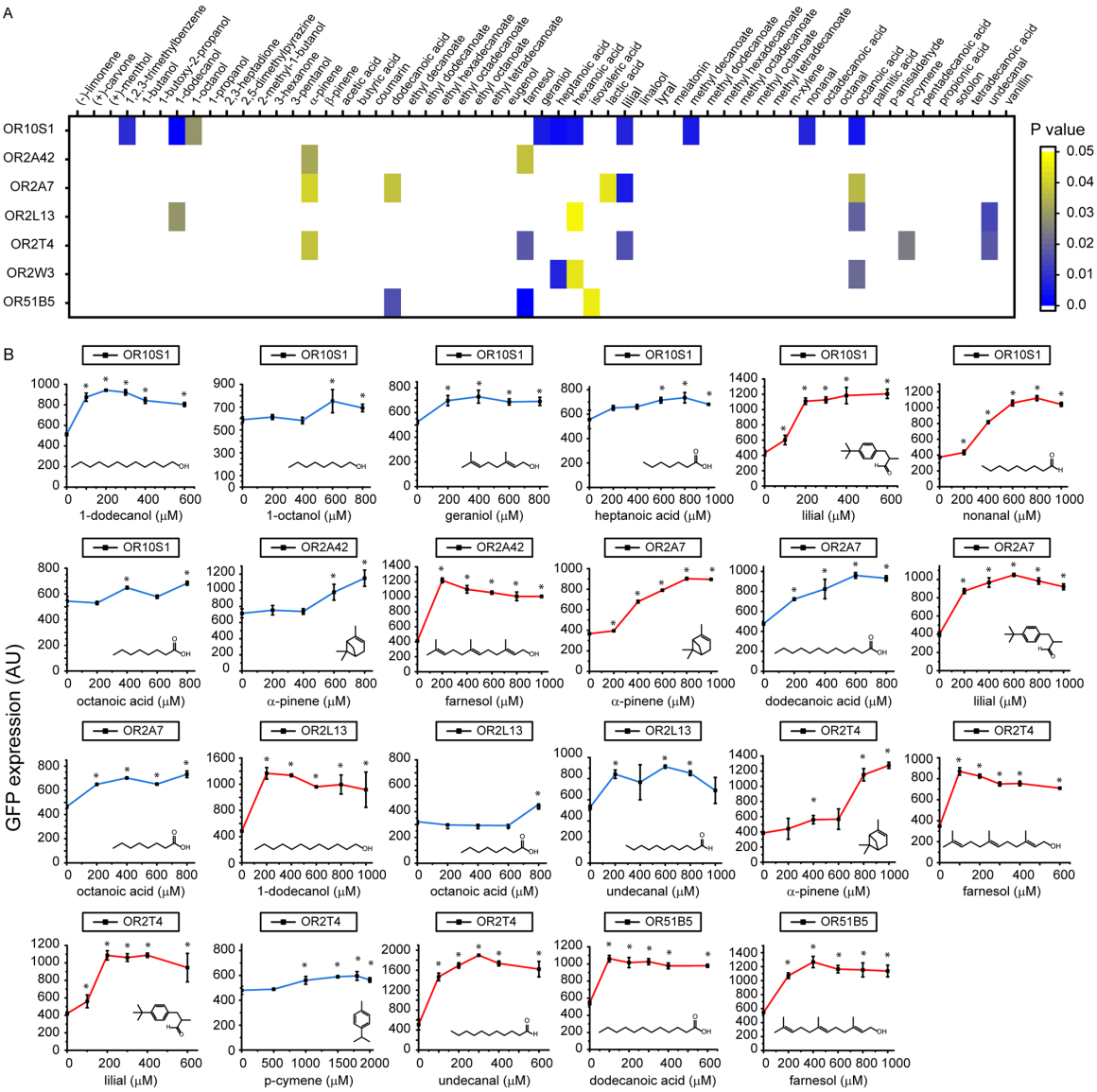

Figure 2.3 - Rapid deorphanization of human olfactory receptors (ORs). (A) Heat map of *P* values of chemicals leading to a statistically significant increase in the level of GFP expression when compared to the dimethyl sulfoxide (DMSO) control (*P* < 0.05). (B) Dose-response curves of ORs with 23 chemical hits. Blue curves are for chemicals leading to a statistically significant increase in the intensity of the signal after activation. Red curves are for chemicals resulting in a >2-fold increase in the intensity of the signal after activation. Asterisks denote statistically significant increases (*P* < 0.05) in the intensity of the signal after activation when compared to no chemical.

identity with the closest sequences being OR2A7 and OR2A42 (71.3%) and the most distant sequences being OR51B5 and OR2L13 (26.2%) (Table 2.1).

2.4.3 Verification of Colon OR Expression in Yeast.

To verify the yeast expression of the seven human ORs, we fused green fluorescent protein (GFP) to the C-terminus of the ORs. The seven ORs are expressed in yeast and could be found at the cell membrane (Figure 2.1). The OR expression pattern was sequence-dependent. While OR2A42 was mostly localized to the cell membrane, OR10S1 was expressed throughout yeast. OR2T4 and OR2W3 showed a punctuated pattern, i.e., the ORs had challenges translocating to the membrane, likely accumulating in the

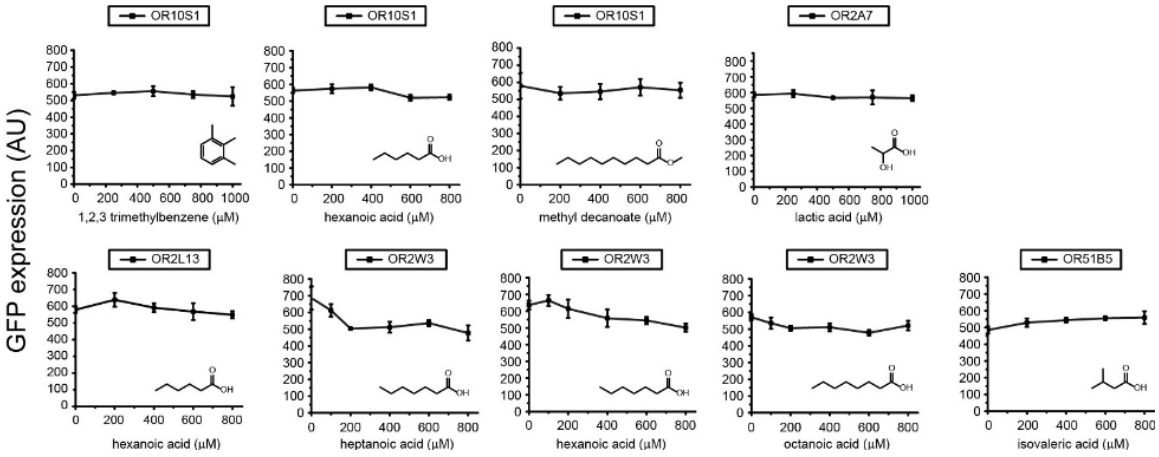

Figure 2.4 – Dose-response curves of the nine OR/chemical pairs that did not result in a statistically larger increase in the intensity of the signal after activation. All measurements were carried out in triplicate, and means ± the standard deviation are shown.

endoplasmic reticulum. Of note, the OR sensor strain amplifies the chemical signal detected by the OR; i.e., activation of the OR leads to the activation of a large number of transcription factors that go on to activate GFP expression. Thus, it is not necessary to have a large number of ORs on the cell surface to detect GFP expression.

2.4.4 Rapid Screening of Colon ORs against the 57-Member Chemical Panel.

Each OR was screened in triplicate against 57 chemicals and DMSO as a control. Chemicals resulting in a statistically significantly increase in the level of GFP expression ($p < 0.05$) when compared to the DMSO control were considered hits (Figure 2.3). There were a total of 32 chemical hits. OR10S1 had the largest number of hits (10). OR2T4 and OR2A7 had five hits each. OR2A42 had the fewest number of chemical hits (two).

2.4.5 Secondary Screening of OR Chemical Hits.

To validate the chemical hits from the rapid screening stage, we determined dose-response curves of the OR/chemical hit pairs (Figure 2.3 and Figure 2.4). Chemical hits were validated if at any of the concentrations tested in the dose-response curve there was a statistically significant increase in the signal after activation when compared to the DMSO control. In the case of OR10S1, seven of 10 chemical hits were validated, with lilial and nonanal leading to a >2-fold increase in the signal after activation. All five OR2T4 chemical hits were validated, with α-pinene, farnesol, lilial, and undecanal showing a >2-fold increase in the signal. The two OR2A42 chemical hits were validated, with farnesol resulting in a 2-fold increase in the signal after activation. Four of the five OR2A7 chemical hits were validated, with α-pinene and lilial leading to a >2-fold increase in the signal after activation. In the case of OR2L13, three of the four chemical hits were validated, with dodecanol and undecanal resulting in a >2-fold increase in the signal. Two of the three OR51B5 chemical hits were validated with dodecanoic and farnesol leading to a >2-fold increase in the signal after activation. None of the three OR2W3 chemical hits were validated. Taken together, we validated at least one hit for each OR except for OR2W3. We find multiple hits for some ORs, which is consistent with olfactory receptors' tendency to bind a family of chemicals.

2.4.6 Confirming OR-Dependent GFP Expression.

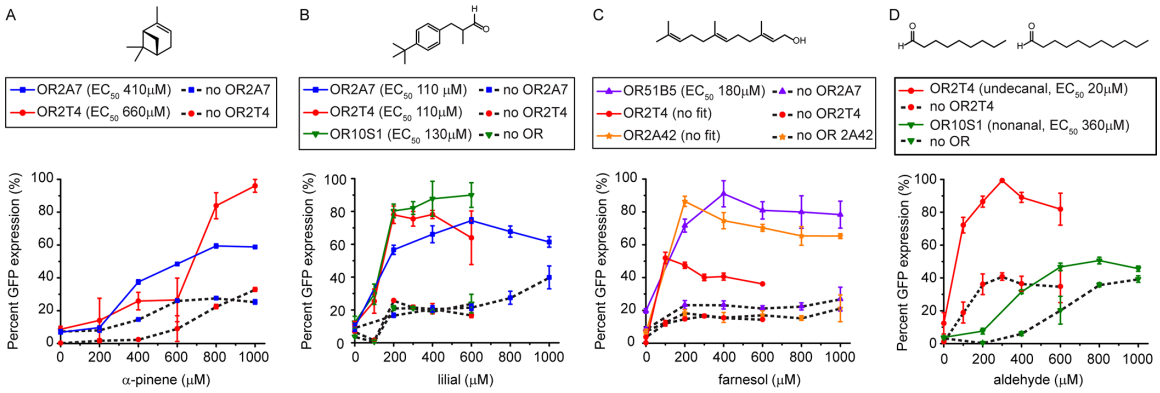

Figure 2.5 – Confirming OR-dependent activation with validated chemicals. Dose-response curves of validated chemicals in the presence (solid lines) and absence (dotted lines) of ORs: (A) OR2A7 and OR2T4 with α-pinene, (B) OR2A7, OR2T4, and OR10S1 with lilial, (C) OR51B5, OR2T4, and OR2A42 with farnesol, and (D) OR2T4 and OR10S1 with undecanal and nonanal, respectively. All measurements were carried out in triplicate, and means ± the standard deviation are shown.

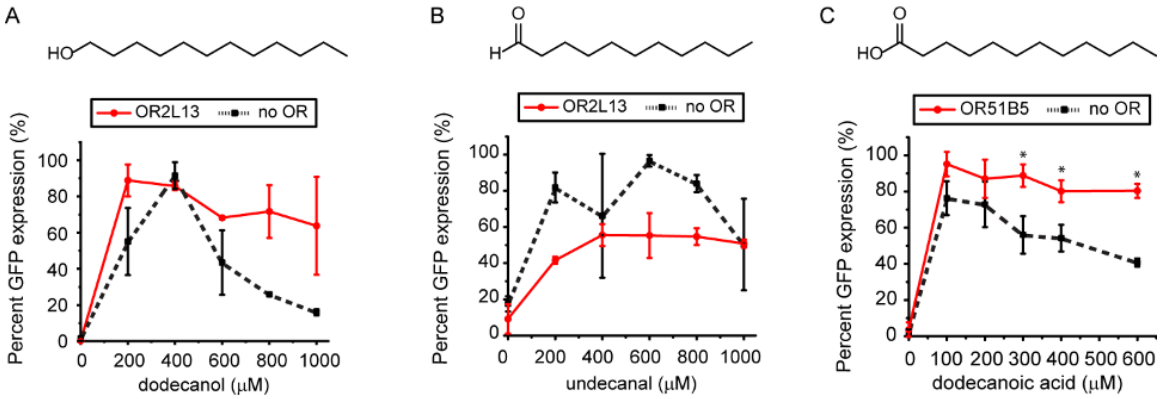

Figure 2.6 – Dose-response curves of three OR/chemical pairs that do not show OR-dependent GFP expression. Dose-response curves of OR2L13 and OR51B5 with validated chemicals. A. OR2L13 with dodecanol. B. OR2L13 with undecanal. C. OR51B5 with dodecanoic acid.

To confirm that the validated chemicals lead to cell fluorescence via OR activation and not an alternative mechanism, we determined dose-response curves of the validated chemicals with a control strain carrying an empty vector in place of the OR and the GFP

reporter plasmid (Figure 2.5 and Figure 2.6). We focused on OR/ chemical pairs resulting in a ≥2-fold increase in the signal after activation: pinene with OR2A7 and OR2T4; lilial with OR2A7, OR2T4, and OR10S1; farnesol with OR51B5, OR2T4, and OR2A52; nonanal with OR10S1; and undecanal with OR2T4. The chemicals inhibit cell growth, but the cells remain at the same optical density as at the start of the experiment (Figure 2.7). We fitted the OR/chemical pair data to a dose-response equation to calculate EC_{50}s. To reliably compare the responses of the chemicals, we ran OR and no OR control experiments pairwise on the same day.

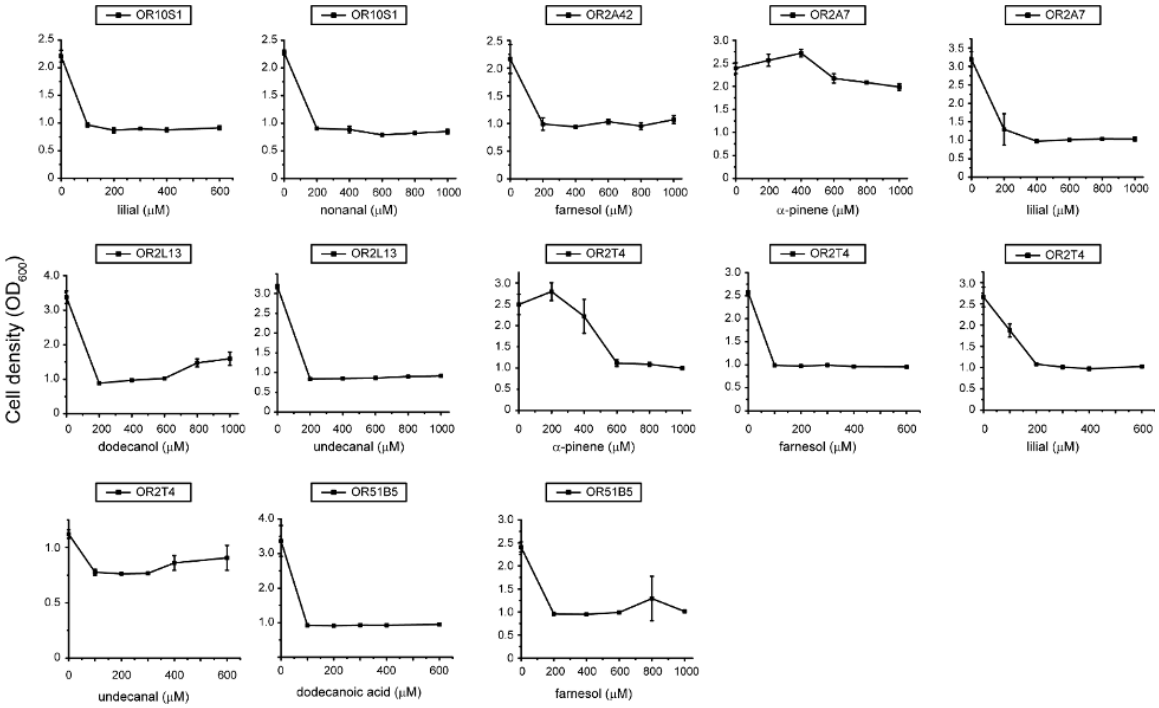

Figure 2.7 – Chemical toxicity to yeast-based OR sensor. Yeast cell density was measured after the 4-hour incubation step in the OR chemical sensing protocol. The no chemical data point has 1% DMSO. While the yeast cells double in the absence of chemical, their cell density is static in the presence of chemical.

Pinene elicits basal GFP expression in the absence of a receptor (Figure 2.5A). In the presence of pinene, the OR2A7 EC_{50} is 412 μM while the OR2T4 EC_{50} is 659 μM. In

the presence of OR2T4, pinene addition results in a 3-fold increase in GFP expression when compared to the no receptor control. In contrast, in the presence of OR2A7, pinene addition results in an only 2.3-fold increase in percent GFP expression when compared to the no receptor control. The maximal level of GFP expression of OR2T4 is 63% higher than that of OR2A7.

Lilial elicits a basal GFP expression that is comparable to that of pinene (Figure 2.5B). The EC_{50}s of OR2A7 and OR2T4 with lilial are almost indistinguishable at 110 and 107 μM, respectively. Lilial shows a lower chemical potency with OR10S1 with an EC_{50} of 129 μM. Lilial elicits GFP expression with OR10S1, resulting in a 3.8-fold increase in percent GFP expression when compared to the no receptor control.

Farnesol results in a basal GFP expression that is comparable to that seen with pinene and lilial (Figure 2.5C). The farnesol response obtained in the presence of OR2T4 and OR2A42 could not be fitted to a dose-response equation with the curves resembling an on/off response. The response of OR51B5 to farnesol could be fitted to a dose-response curve, resulting in an EC_{50} of 181 μM, and it has a 3-fold increase in percent GFP expression compared to the no receptor control.

The two aldehydes, undecanal and nonanal, elicit a slightly higher basal GFP expression than pinene, lilial, or farnesol (Figure 2.5D). Although OR10S1 shows an increase in its signal after activation upon nonanal addition, the no OR control shows a similar increase in GFP expression. In the presence of OR2T4, the addition of undecanal results in a 2.3-fold increase in percent GFP expression when compared to the no receptor control.

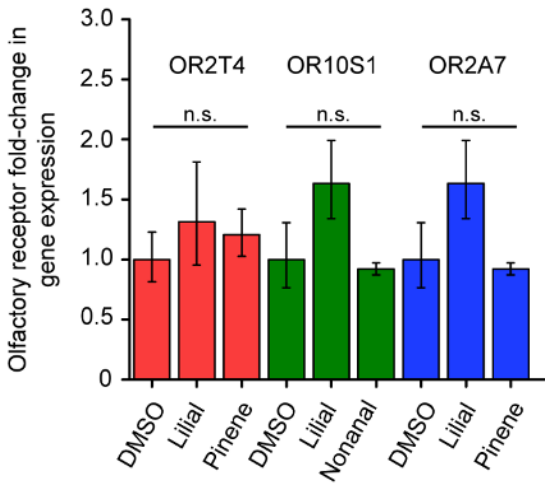

Figure 2.8 – Changes in olfactory receptor gene expression levels in the presence of chemicals. Red bars: OR2T4. Green bars: OR10S1. Blue bars: OR2A7. The gene expression experiments were run 600 μM lilial and nonanal, and 1000 μM pinene. DMSO controls were compared to OR/chemical pairs for OR2T4, OR10S1, and OR2A7 and in all cases $P > 0.05$.

Taken together, pinene is a confirmed ligand for OR2A7 and OR2T4, lilial is a confirmed ligand for OR2A7, OR2T4, and OR10S1, while undecanal is a confirmed ligand for OR2T4. Addition of these chemicals does not elicit an increase in the level of OR gene expression when compared to the DMSO control (Figure 2.8). Thus, the increase in GFP expression is due to signal transfer and not an increase in the number of olfactory receptors expressed. Farnesol does not result in a dose-response fit with OR2T4 or OR2A42; thus, it is not a ligand for these ORs. Although addition of farnesol to OR51B5 did result in a dose-response fit, the response has an overall on/off behavior (except for the data point at 200 μM farnesol). To test if farnesol generally increases the level of GFP expression in the presence of ORs, we measured the response of OR2W3 and OR2A7 upon addition of farnesol (Figure 2.9). In the presence of OR2W3 or OR2A7, farnesol elicits GFP expression, yet the data do not fit a dose-response curve. In conclusion, farnesol nonspecifically activates GFP expression in the OR-based sensors.

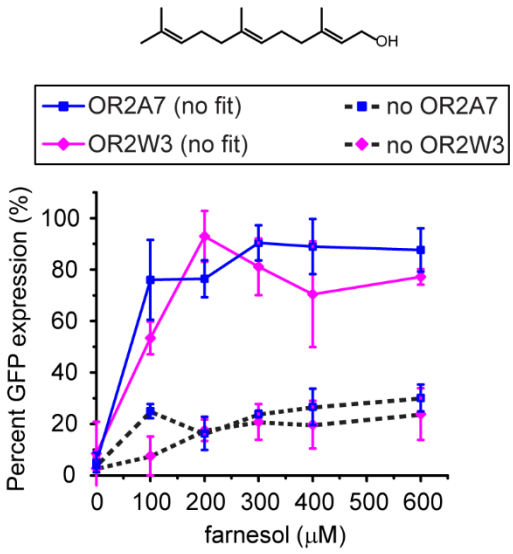

Figure 2.9 – Dose-response curves of OR2A7 and OR2W3 with farnesol.

2.4.7 Understanding the Chemical Activation Profile of OR2A7 and OR2T4.

We determined the dose-response curves of OR2A7 and OR2T4 with chemicals that have stereochemistry and substructure different from those of the identified hits, pinene and lilial. Activation of the ORs is dependent on pinene stereochemistry. OR2A7 shows an 18% weaker response with *β*-pinene than *α*-pinene, while the response of OR2T4 to *β*-pinene is similar to the GFP expression of the no OR control (Figure 2.10A,B). Both OR2T4 and OR2A7 are activated by lilial. We probed the activation profile of OR2T4 and OR2A7 with 3-phenylbutyraldehyde (3PB), which retains the aldehyde found in lilial but lacks the *tert*-butyl group, and *tert*-butylbenzene (TBB), which lacks the aldehyde moiety but retains the phenyl and *tert*-butyl groups. We find that to activate OR2T4 and OR2A7, the *tert*-butyl group and the aldehyde side chain are necessary (Figure 2.10C,D). Although TBB shows an increase in the level of GFP expression at >600 μM, the no receptor control shows a similar increase in the level of GFP expression, making the signal observed not OR dependent.

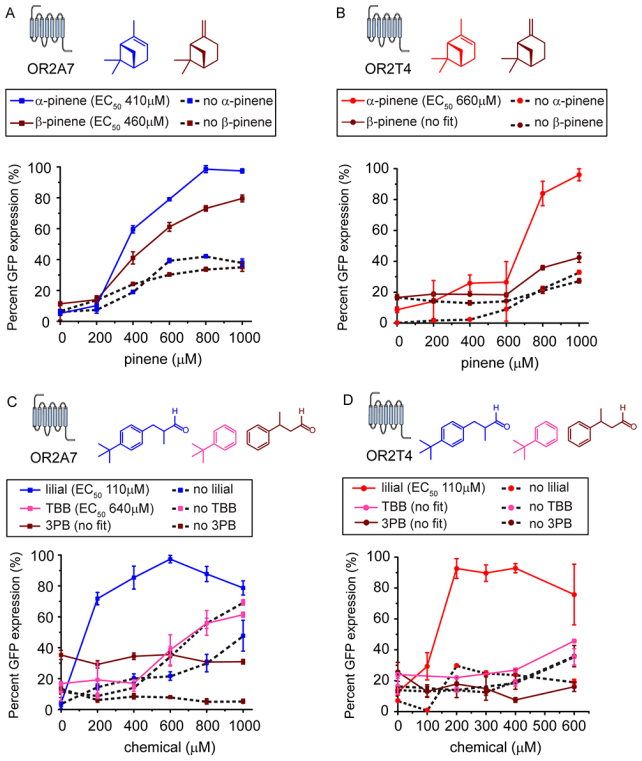

Figure 2.10 – Chemical activation profiles of OR2A7 and OR2T4. Dose-response curves of (A) OR2A7 with *α*-pinene and *β*-pinene, (B) OR2T4 with *α*-pinene and *β*-pinene, (C) OR2A7 with lilial, 3-phenylbutyraldehyde (3PB), and *tert*-butylbenzene (TBB), and (D) OR2T4 with lilial, 3PB, and TBB. All measurements were carried out in triplicate and means ± the standard deviation are shown.

Table 2.2 – Goodness of fit (R^2) for dose-response curves.

Figure	OR	Chemical	R^2
Figure 2.5A, Figure 2.10A	OR2A7	*α*-pinene	0.98946
Figure 2.5A, Figure 2.10B	OR2T4	*α*-pinene	0.99269
Figure 2.5B, Figure 2.10C	OR2A7	lilial	0.97004
Figure 2.5B, Figure 2.10D	OR2T4	lilial	0.99854
Figure 2.5B	OR10S1	lilial	0.99623
Figure 2.5C	OR51B5	farnesol	0.99585
Figure 2.5D	OR2T4	undecanal	0.90762
Figure 2.5D	OR10S1	nonanal	0.99651
Figure 2.10A	OR2A7	*β*-pinene	0.99659
Figure 2.10C	OR2A7	*tert*-butylbenzene	0.99853

2.5 Conclusion

Here, we leveraged a previously engineered GPCR-based yeast sensing strain to generate sensors for seven ORs found in the human colon. Yeast's robustness and rapid doubling time allowed us to quickly screen each of the seven colon ORs against 57 chemicals and deorphanize two receptors, OR2T4 (α-pinene, lilial, and undecanal) and OR10S1 (lilial), and identify two new inexpensive ligands for OR2A7 (α-pinene and lilial). The rapid deorphanization workflow can be repeated to deorphanize other ORs and can be used, in the future, to identify the endogenous ligands of OR2T4, OR10S1, and OR2A7 in the colon. The yeast-based sensor used in this work links the human ORs to GFP expression via yeast G_α subunit GPA1. Sometimes, an OR is activated by different ligands depending on the G_α subunit to which they couple.[24,25] Sometimes, ORs are activated by a ligand independent of the G_α subunit to which it couples, and using G_{olf} enhances the signal.[26,27] In the future, the OR/ligand pairs identified in this work can be coupled to G_{olf} to determine the situation under which they fall. This can be accomplished using the mammalian OR sensor system that expresses G_{olf}[5,28] or by expressing G_{olf} in yeast.

We did not find ligands for OR2A42, OR2L13, OR2W3, or OR51B5. Although these ORs are expressed in yeast, it is possible that they are not coupling to the yeast machinery. Alternatively, ligands for these receptors may not be present among the 57 chemicals tested. Deorphanization of these ORs will likely require the use of G_{olf}/GPA1 fusion protein for improved coupling to the yeast machinery and a larger chemical library against which to screen.

2.6 References

[1] Spehr, M., and Munger, S. D. (2009) Olfactory receptors: G protein-coupled receptors and beyond. *J. Neurochem. 109,* 1570−1583.

[2] Buck, L. B. (2004) Olfactory receptors and odor coding in mammals. *Nutr. Rev. 62,* S184−S188.

[3] Peterlin, Z., Firestein, S., and Rogers, M. E. (2014) The state of the art of odorant receptor deorphanization: A report from the orphanage. *J. Gen. Physiol. 143,* 527−542.

[4] de March, C. A., Ryu, S., Sicard, G., Moon, C., and Golebiowski, J. (2015) Structure−odour relationships reviewed in the postgenomic era. *Flavour Fragrance J. 30,* 342−361.

[5] Mainland, J. D., Keller, A., Li, Y. R., Zhou, T., Trimmer, C., Snyder, L. L., Moberly, A. H., Adipietro, K. A., Liu, W. L., Zhuang, H., Zhan, S., Lee, S. S., Lin, A., and Matsunami, H. (2014) The missense of smell: functional variability in the human odorant receptor repertoire. *Nat. Neurosci. 17,* 114−120.

[6] Mainland, J. D., Li, Y. R., Zhou, T., Liu, W. L. L., and Matsunami, H. (2015) Human olfactory receptor responses to odorants. *Sci. Data 2,* 150002.

[7] Bushdid, C., de March, C. A., Fiorucci, S., Matsunami, H., and Golebiowski, J. (2018) Agonists of G-Protein-Coupled Odorant Receptors Are Predicted from Chemical Features. *J. Phys. Chem. Lett. 9,* 2235−2240.

[8] Flegel, C., Manteniotis, S., Osthold, S., Hatt, H., and Gisselmann, G. (2013) Expression profile of ectopic olfactory receptors determined by deep sequencing. *PLoS One 8,* No. e55368.

[9] Feldmesser, E., Olender, T., Khen, M., Yanai, I., Ophir, R., and Lancet, D. (2006) Widespread ectopic expression of olfactory receptor genes. BMC Genomics 7, 121.

[10] Spehr, M., Gisselmann, G., Poplawski, A., Riffell, J. A., Wetzel, C. H., Zimmer, R. K., and Hatt, H. (2003) Identification of a testicular odorant receptor mediating human sperm chemotaxis. Science 299, 2054−2058.

[11] Pluznick, J. L., Protzko, R. J., Gevorgyan, H., Peterlin, Z., Sipos, A., Han, J., Brunet, I., Wan, L. X., Rey, F., Wang, T., Firestein, S. J., Yanagisawa, M., Gordon, J. I., Eichmann, A., Peti-Peterdi, J., and Caplan, M. J. (2013) Olfactory receptor responding to gut microbiota-derived signals plays a role in renin secretion and blood pressure regulation. *Proc. Natl. Acad. Sci. U. S. A. 110,* 4410−4415.

[12] Leja, J., Essaghir, A., Essand, M., Wester, K., Oberg, K., Tötterman, T. H., Lloyd, R., Vasmatzis, G., Demoulin, J. B., and Giandomenico, V. (2009) Novel markers for enterochromaffin cells and gastrointestinal neuroendocrine carcinomas. *Mod. Pathol. 22,* 261− 272.

[13] Priori, D., Colombo, M., Clavenzani, P., Jansman, A. J., Lalleѕ, J. P., Trevisi, P., and Bosi, P. (2015) The Olfactory Receptor OR51E1 Is Present along the Gastrointestinal Tract of Pigs, Co-Localizes with Enteroendocrine Cells and Is Modulated by Intestinal Microbiota. *PLoS One 10,* No. e0129501.

[14] Fleischer, J., Bumbalo, R., Bautze, V., Strotmann, J., and Breer, H. (2015) Expression of odorant receptor Olfr78 in enteroendocrine cells of the colon. *Cell Tissue Res. 361,* 697−710.

[15] Zhang, X., Bedigian, A. V., Wang, W., and Eggert, U. S. (2012) G protein-coupled receptors participate in cytokinesis. *Cytoskeleton 69,* 810−818.

[16] Veitinger, T., Riffell, J. R., Veitinger, S., Nascimento, J. M., Triller, A., Chandsawangbhuwana, C., Schwane, K., Geerts, A., Wunder, F., Berns, M. W., Neuhaus, E. M., Zimmer, R. K., Spehr, M., and Hatt, H. (2011) Chemosensory Ca2+ Dynamics Correlate with Diverse Behavioral Phenotypes in Human Sperm. *J. Biol. Chem. 286,* 17311− 17325.

[17] Manteniotis, S., Wojcik, S., Göthert, J. R., Dürig, J., Dührsen, U., Gisselmann, G., and Hatt, H. (2016) Deorphanization and characterization of the ectopically expressed olfactory receptor OR51B5 in myelogenous leukemia cells. *Cell Death Discovery 2,* 16010.

[18] Flegel, C., Vogel, F., Hofreuter, A., Schreiner, B. S., Osthold, S., Veitinger, S., Becker, C., Brockmeyer, N. H., Muschol, M., Wennemuth, G., Altmüller, J., Hatt, H., and Gisselmann, G. (2016) Characterization of the Olfactory Receptors Expressed in Human Spermatozoa. *Front. Mol. Biosci. 2,* 73.

[19] Fujita, Y., Takahashi, T., Suzuki, A., Kawashima, K., Nara, F., and Koishi, R. (2007) Deorphanization of Dresden G protein-coupled receptor for an odorant receptor. *J. Recept. Signal Transduction Res. 27,* 323−334.

[20] Tsai, T., Veitinger, S., Peek, I., Busse, D., Eckardt, J., Vladimirova, D., Jovancevic, N., Wojcik, S., Gisselmann, G., Altmüller, J., Staň der, S., Luger, T., Paus, R., Cheret, J., and Hatt, H. (2017) Two olfactory receptors-OR2A4/7 and OR51B5-differentially affect epidermal proliferation and differentiation. *Exp. Dermatol. 26,* 58−65.

[21] Mukherjee, K., Bhattacharyya, S., and Peralta-Yahya, P. (2015) GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synth. Biol. 4,* 1261−1269.

[22] Wenderski, T. A., Stratton, C. F., Bauer, R. A., Kopp, F., and Tan, D. S. (2015) Principal Component Analysis as a Tool for Library Design: A Case Study Investigating Natural Products, Brand-Name Drugs, Natural Product-Like Libraries, and Drug-Like Libraries. *Methods Mol. Biol. 1263,* 225−242.

[23] Zhang, X., De la Cruz, O., Pinto, J. M., Nicolae, D., Firestein, S., and Gilad, Y. (2007) Characterizing the expression of the human olfactory receptor gene family using a novel DNA microarray. *Genome Biol. 8,* R86.

[24] Shirokova, E., Schmiedeberg, K., Bedner, P., Niessen, H., Willecke, K., Raguse, J. D., Meyerhof, W., and Krautwurst, D. (2005) Identification of specific ligands for orphan olfactory receptors. G protein-dependent agonism and antagonism of odorants. *J. Biol. Chem. 280,* 11807−11815.

[25] Ukhanov, K., Bobkov, Y., Corey, E. A., and Ache, B. W. (2014) Ligand-selective activation of heterologously-expressed mammalian olfactory receptor. *Cell Calcium 56,* 245−256.

[26] Minic, J., Persuy, M. A., Godel, E., Aioun, J., Connerton, I., SAlesse, R., and Pajot-Augy, E. (2005) Functional expression of olfactory receptors in yeast and development of a bioassay for odorant screening. *FEBS J. 272,* 524−537.

[27] Fukutani, Y., Nakamura, T., Yorozu, M., Ishii, J., Kondo, A., and Yohda, M. (2012) The N-terminal replacement of an olfactory receptor fort he development of a yeast-based biomimetic odor sensor. *Biotechnol. Bioeng. 109,* 205−212.

[28] Zhuang, H., and Matsunami, H. (2008) Evaluating cell-surface expression and measuring activation of mammalian odorant receptors in heterologous cells. *Nat. Protoc. 3,* 1402−1413.

CHAPTER 3. IDENTIFICATION OF THREE ANTIMICROBIALS ACTIVATING SEROTONIN RECEPTOR 4 IN COLON CELLS

Reprinted with permission from:

Yasi, E. A., Allen, A. A., Sugianto, W. and Peralta-Yahya, P. *ACS Synth. Biol.* **2019,** *8 (12)* 2710-2717.

3.1 Abstract

The serotonin receptor 4b (5-HTR_{4b}) is expressed throughout the gastrointestinal tract, and its agonists are used in the treatment of irritable bowel syndrome with constipation (IBS-C). Today, there are no rapid assays for the identification of 5-HTR_{4b} agonists. Here, we developed a luciferase-based 5-HTR_{4b} assay capable of assessing one compound per second with a 38-fold dynamic range and nM limit of detection for serotonin. We used the assay to screen more than 1000 natural products and anti-infection agents and identified five new 5-HTR_{4b} ligands: hordenine, halofuginone, proflavine, ethacridine, and revaprazan. We demonstrate that hordenine (antibiofilm), halofuginone (antiparasitic), and revaprazan (gastric acid reducer) activate 5-HTR_{4b} in human colon epithelial cells, leading to increased cell motility or wound healing. The 5-HTR_{4b} assay can be used to screen larger pharmaceutical libraries to identify novel treatments for IBS-C. This work shows that antimicrobials interact not only with the gut microbiota, but also with the human host.

3.2 Introduction

In humans, 95% of serotonin (5-HT) is found in the gastrointestinal (GI) tract,[1] where release and reception of 5-HT transmits information from the gut lumen to gut nerve cells and smooth muscles. Of the seven 5-HT receptor families, 5-HTR_4 is broadly expressed in the gut: on nitrergic neurons that control smooth muscle relaxation, cholinergic and nitrergic neurons that control muscle contraction and relaxation, enterocytes that control chemical transport, and enteroendocrine cells that control the secretion of gastrointestinal hormones.[2,3] Specifically, 5-HTR_{4b} is highly expressed in the jejunum, ileum, and colon.[4] 5-HTR_4 has been implicated in irritable bowel syndrome (IBS), which affects 15% of the world population.[4,5] Agonists of 5-HTR_4 are used for the treatment of irritable bowel syndrome with constipation (IBS-C), relieving constipation, abdominal pain, and bloating.

One of the major challenges in identifying novel 5-HTR_4 agonists is the dearth of 5-HTR high-throughput assays to rapidly assess large libraries of chemicals. The two-day culture time required to test cell motility using colon cells, which naturally express 5-HTR_4, would be time-prohibitive for a primary screening tool. Commercial G-protein coupled receptor (GPCR)-based assays, such as SelectScreen (Thermo Fisher) and gpcrMAX^{SM} (Eurofins) lack screens for any member of the 5-HTR_4 family. To our knowledge, only one large-scale 5-HTR_4 screen has been performed to date, against 976 ToxCast chemicals using guinea pig brain tissue and radio-labeled ligand displacement.[6]

Previously, we engineered a fluorescence-based 5-HTR_{4b} assay in yeast by linking expression of human 5-HTR_{4b} on the yeast surface to green fluorescent protein (GFP) expression.[7] The fluorescent reporter, however, posed some practical limitations.

Fluorescent readout requires a 4 h chemical incubation step for robust GFP expression, and a high-throughput flow cytometer for signal readout, thus limiting the assay throughput to one 96-well plate per hour. Additionally, the GFP-based 5-HTR$_4$ assay has only a 3-fold dynamic range and one order of magnitude linear range after activation with serotonin.

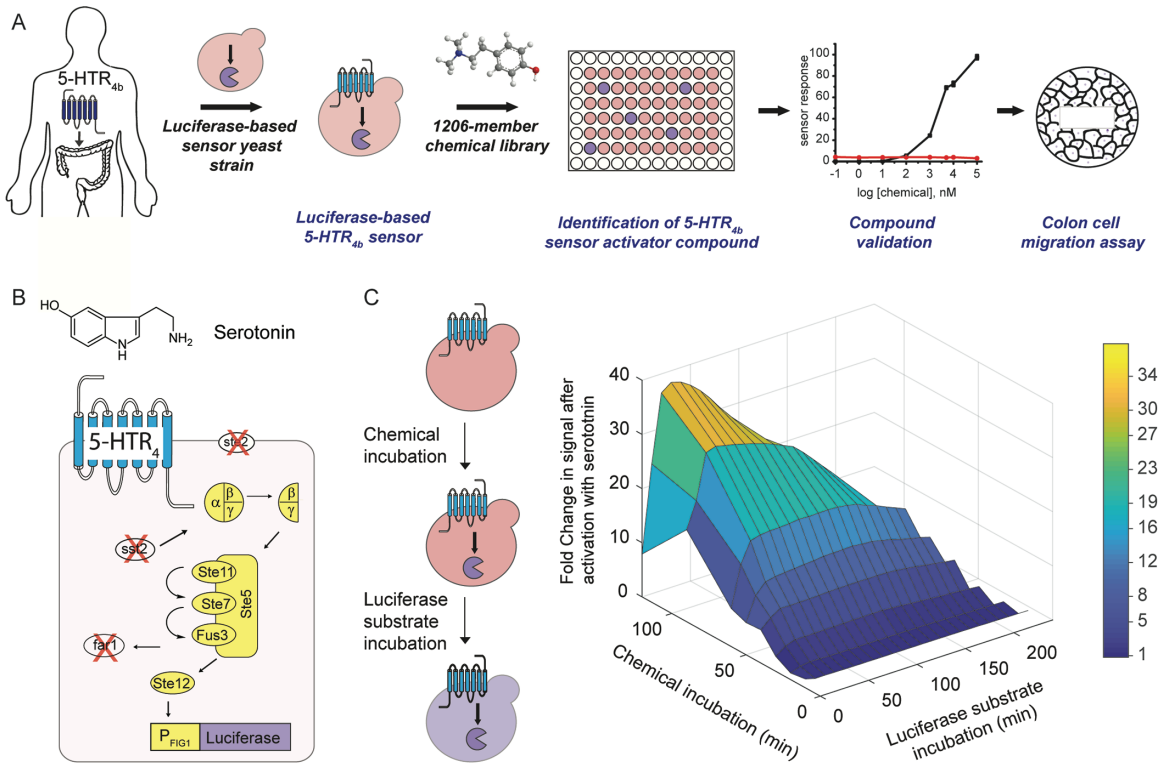

Figure 3.1 – Luciferase-based 5-HTR$_4$ assay development. (A) Workflow for the identification and validation of HTR$_{4b}$ agonists. (B) Luciferase-based 5-HTR$_4$ assay: human 5-HTR$_{4b}$ was expressed on the cell surface of a yeast engineered to link receptor activation to reporter to luciferase gene expression via the yeast mating pathway. (C) 5-HTR$_4$-assay optimization at a pH = 7 and 100 mg/L serotonin.

Here, we developed a luciferase-based 5-HTR$_{4b}$ assay with an overall assay time of 2.5 h, enabling the use of a luminescent plate reader, and achieving a screening throughput of one compound per second. Next, we used the assay to screen 1206 compounds coming from two commercial chemical libraries, a natural products library and an anti-infection

library, and discovered five previously unidentified 5-HTR_{4b} ligands. We validated three of the five ligands, hordenine, halofuginone, and revaprazan, as 5-HTR_{4b} agonists in mammalian cells as they increase motility or wound healing in human colon epithelial cells. (Figure 3.1A).

This work has three significant outcomes. First, the same compound discovery workflow can be used, in the future, to screen pharmaceutical libraries for the identification of novel 5-HTR_{4b} agonists for the treatment of IBS-C. Second, the increased assay signal provided by the luciferase reporter, when compared to the GFP reporter, should enable the generation of other high-throughput GPCR-based assays by simply swapping the receptor from the cell surface. Third, as antibiofilm (hordenine) and antiparasitic (halofuginone) agents affect colon cell motility and/or wound healing, antimicrobials may interact not only with the gut microbiota, but also with the human host leading to potential changes in gut movement and secretion. Finally, with a 5-HTR_{4b} high-throughput assay in hand, we can now screen gut microbiota metabolites to further understand the link between host and gut microbiome.

3.3 Results

3.3.1 Luciferase-Based 5-HTR_{4b} Assay Development.

We replaced the GFP reporter from our GFP-based 5-HTR_{4b} assay[7] with NanoLuc luciferase,[8] which we optimized for yeast expression (Figure 3.1B). As the pH of the GI tract hovers between 5.7 and 7.4, we optimized the assay at a pH of 7. The 5-HTR_{4b} assay is composed of two steps, (i) ligand incubation leading to luciferase expression, and (ii)

luciferase substrate incubation leading to luminescence. By co-optimizing both incubation steps using serotonin, we concluded that a 2 h ligand incubation followed by a 30 min luciferase substrate incubation results in the fastest overall assay conditions (2.5 h) with the highest signal increase after activation (38-fold) (Figure 3.1C). Of note, multiple 96-well plates can be incubated simultaneously, as reading the plate for luminescence takes 2 min.

3.3.2 Luciferase-Based 5-HTR$_{4b}$ Assay Validation.

We demonstrate that the assay detects three known 5-HTR$_4$ agonists for the treatment of IBS-C: tegaserod, prucalopride, and mosapride,[9,10] and four other agonists used to treat gastroesophageal reflux (cisapride),[11] depression (RS67333),[12] anxiety (zacopride),[13] and nausea (metoclopramide)[10,14] (Figure 3.2A−H). To verify that the agonists led to cell luminescence due to 5-HTR$_{4b}$ activation and not via an alternative mechanism, we performed dose-response curves with the agonists using a control strain carrying the luciferase reporter plasmid and a blank plasmid in lieu of 5-HTR$_{4b}$. On the basis of EC$_{50}$s, the potency of the agonists toward 5-HTR$_{4b}$ are tegaserod (0.3 nM) > RS67333 (11.0 nM) > prucalopride (41.0 nM) > cisapride (69.9 nM) > serotonin (155.0 nM) > mosapride (256.4nM) > zacopride (616.0 nM) > metoclopramide (7.4 μM). These results agree with previous studies that identified tegaserod and RS67333 to be more potent than serotonin,[15,16] and zacopride and metoclopramide to be less potent.[17−19] Taken together, the luciferase-based HTR$_{4b}$ assay is capable of identifying drugs with EC$_{50}$s ranging from the low nM to the μM level.

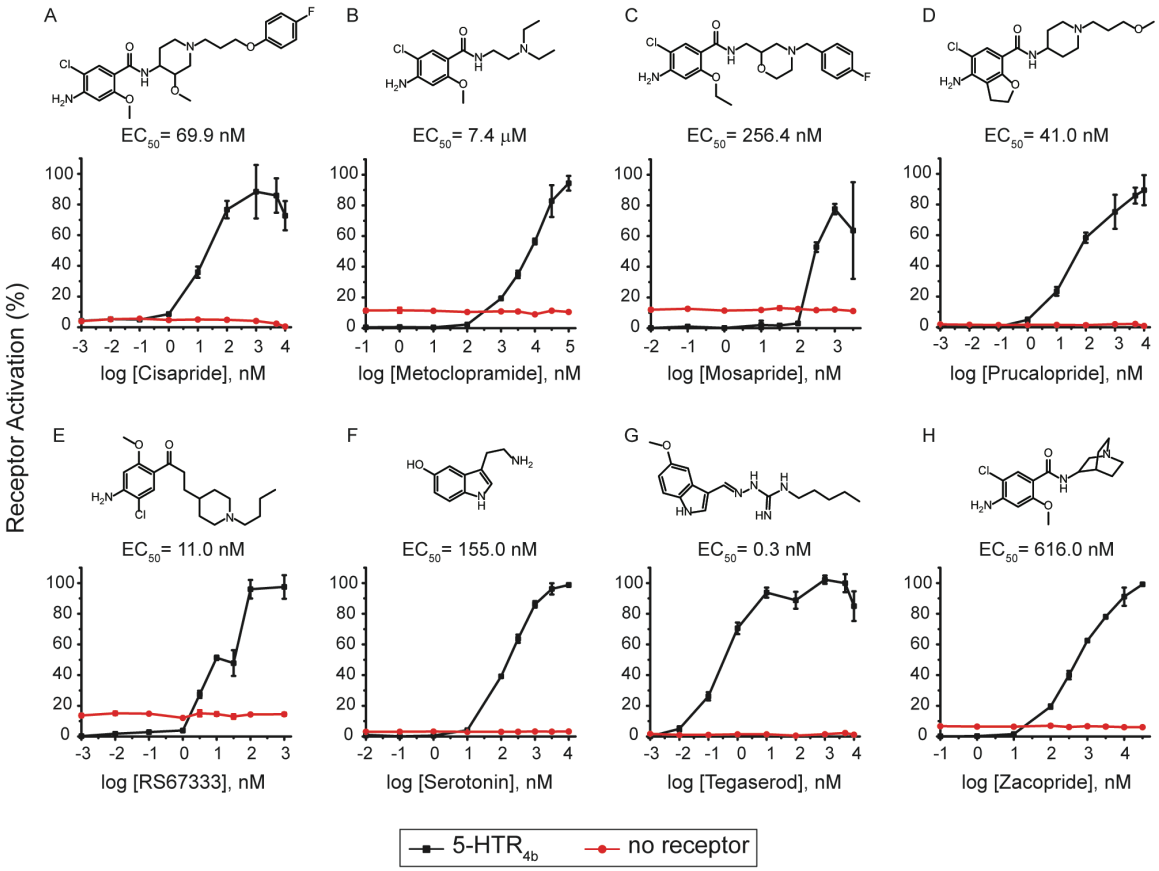

Figure 3.2 – Validation of the 5-HTR$_4$ assay. 5-HTR$_4$ assay dose-response curves with known 5-HTR$_4$ agonists: (A) cisapride, (B) metoclopramide, (C) mosapride, (D) prucalopride, (E) RS-67333, (F) serotonin, (G) tegaserod, (H) zacopride. Data was collected in triplicate. Shown are means ± s.d.

3.3.3 5-HTR$_{4b}$ Assay High-Throughput Screening Validation.

We validated the assay for 96-well plate high-throughput screening using a 3-day plate uniformity experiment[20] (Figure 3.3). The assay had an average *Z* factor of 0.74 and an average coefficient of variation of 7.7%, meeting the two statistical parameters for high-throughput assay acceptance, *i.e.,* a *Z* factor of >0.5[21] and a CV < 10%.[22]

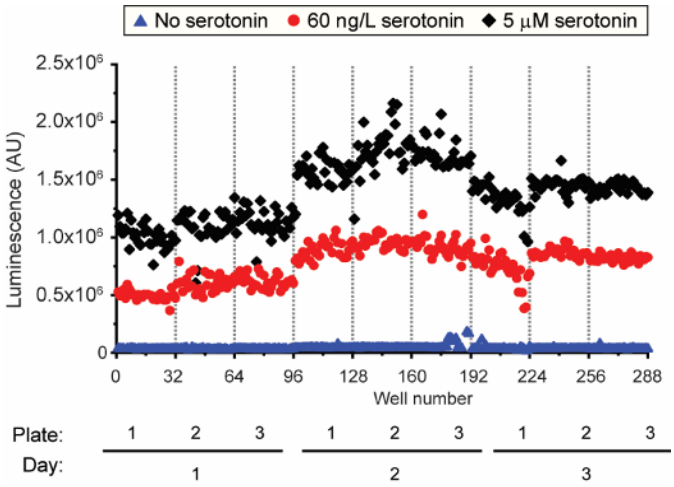

Figure 3.3 – Three-day 96 well-plate validation assay. Luminescence data with no, medium and high levels of serotonin. No obvious intra-plate edge effects or drift observed. The signal activation levels did fluctuate from day to day, but all raw midpoint CVs were below 20%.

3.3.4 Identifying Novel 5-HTR$_{4b}$ Ligands.

GPCRs expressed in the GI tract tend to bind gut microbial metabolites. For example, GPR41, GPR43, and GPR109 bind microbiota produced short chain fatty acids.[23–25] We hypothesize that 5-HTR$_{4b}$ may bind microbial natural products. To explore the range of biological compounds that 5-HTR$_{4b}$ may bind, we used the assay to screen a commercial 803-member chemical natural products library. First, to understand the chemical diversity of the library, we broke down each library member into 23 chemical descriptors to perform a principal component analysis (Figure 3.4A, Table 3.1). The principal components (PC axes) reflect the common and unique variances of the chemical descriptors with the top three PCs accounting for 64% of the cumulative variances, with PC1, PC2, and PC3 capturing 45%, 11%, and 8% of the total variance, respectively. On the basis of their PC scores, chemicals can be separated into eight chemical spaces (CS)

(Table 3.2). All chemical spaces are populated by compounds in the natural products library,

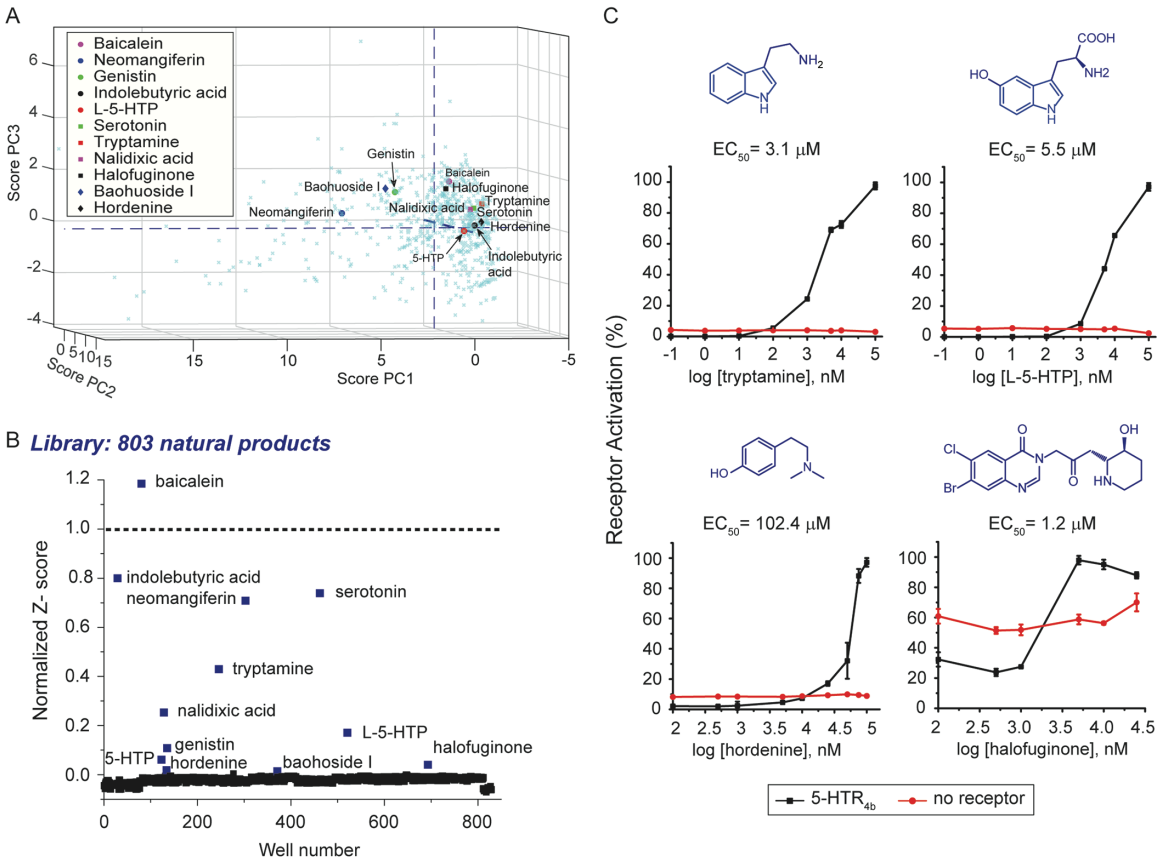

Figure 3.4 – 5-HTR$_4$ assay identifies natural products as novel ligands. (A) Principal component analysis of natural products library with 5-HTR$_{4b}$ screening hits highlighted. (B) 803-member natural product library screening results. The compounds were screened using the 5-HTR$_{4b}$ assay in singlets. *Z*-scores were normalized to the serotonin positive control, which was set to 1 (dotted line). (C) Dose-response curves of the four validated natural product hits. Data was collected in triplicate. Shown are means ± s.d.

Table 3.1 – Table of chemical descriptors used in PCA analysis

Descriptors	Abbreviations	Descriptors	Abbreviations
Molecular Weight	MW	Ring system count	RSC
Carbon atom count	CAC	Size of largest ring	SLR

Hydrogen atom count	HAC	Van der Waals surface area	VSA
Oxygen atom count	OAC	Amine	-
Nitrogen atom count	NAC	Alcohol	-
Hydrogen-bond donors	HBD	Ether	-
Hydrogen-bond acceptor	HBA	Aldehyde	-
Rotatable bonds	RB	Ketone	-
Stereocenter count	SC	Carboxylic Acid	-
Topological polar surface area	TPSA	Ester	-
Number of rings	NOR	Amide	-
Aromatic ring count	ARC		

Table 3.2 – Chemical space composition of the natural product library.

Chemical Space	PC1	PC2	PC3	Members	Population (%)	Chemical characteristics
1	-	-	-	128	16	Mix of aromatics and groups containing N and O atoms
2	+	-	-	88	11	Steric, heterocyclic with O-containing functional groups
3	-	-	+	177	22	small heteroaromatics (fused rings) with O-containing functional groups
4	+	-	+	90	11	Adjacent to CS 3, hence similar trend, except higher C-atom count
5	+	+	-	47	6	chiral with multiple functional groups (phosphate, sulfate, N and O-containing functional groups)
6	-	+	-	124	16	Two distinct groups: A) long C-tail containing only N or O-functional group. B) complex heterocyclic (individual benzene rings) having both N and O-containing functional groups
7	+	+	+	49	6	hydrocarbons and cyclic rings

8	-	+	+	98	12	Adjacent to CS 6, hence similar trend, except heteroaromatics are fused rings
			Total:	801	100	

indicating its chemical diversity. CS3, composed of small heteroaromatics with O-containing functionality, is the highest populated CS, containing 22% of all chemicals in the library. The least populated CSs are CS5 and CS7, each containing 6% of all the chemicals tested. In the future, the chemical diversity of the library could be improved by introducing more compounds with chiral centers and cyclic rings.

The screening of the natural products library resulted in 12 hits: serotonin, tryptamine, 3-indolebutyric acid, L-5-hydroxy-tryptophan (L-5-HTP), D/L-5-hydroxytryptophan (5-HTP), hordenine, genistin, neomangiferin, baohuoside I, baicalein, halofuginone, and nalidixic acid (Figure 3.4B). Previously, tryptamine[26] and 5-HTP[27] have been shown to bind 5-HTR$_4$, and baicalein[28] has been shown to bind 5-HTR$_7$. Hordenine, nalidixic acid, halofuginone, 3-indolebutyric acid, genistin neomangiferin, and baohuoside I have not been previously shown to bind 5-HTR$_4$.[29]

To confirm the natural products library hits, we ran dose-response curves using the 5-HTR$_{4b}$ assay strain, and a control strain carrying the luciferase reporter and a blank plasmid in lieu of 5-HTR$_{4b}$. We confirmed tryptamine and 5-HTP to be ligands of 5-HTR$_4$. We could not confirm 3-indolebutyric acid, nalidixic acid, baicalein, genistin, baohuoside I, and neomangiferin as ligands of 5-HTR$_{4b}$ (Figure 3.5). We find, for the first time, that hordenine (EC_{50} 102.4 μM) and halofuginone (EC_{50} 1.2 μM) are HTR$_{4b}$ ligands (Figure

3.4C). Except for L-5-HTP, all natural products hits, i.e., hordenine, tryptamine, serotonin, and halofuginone, belong to CS8.

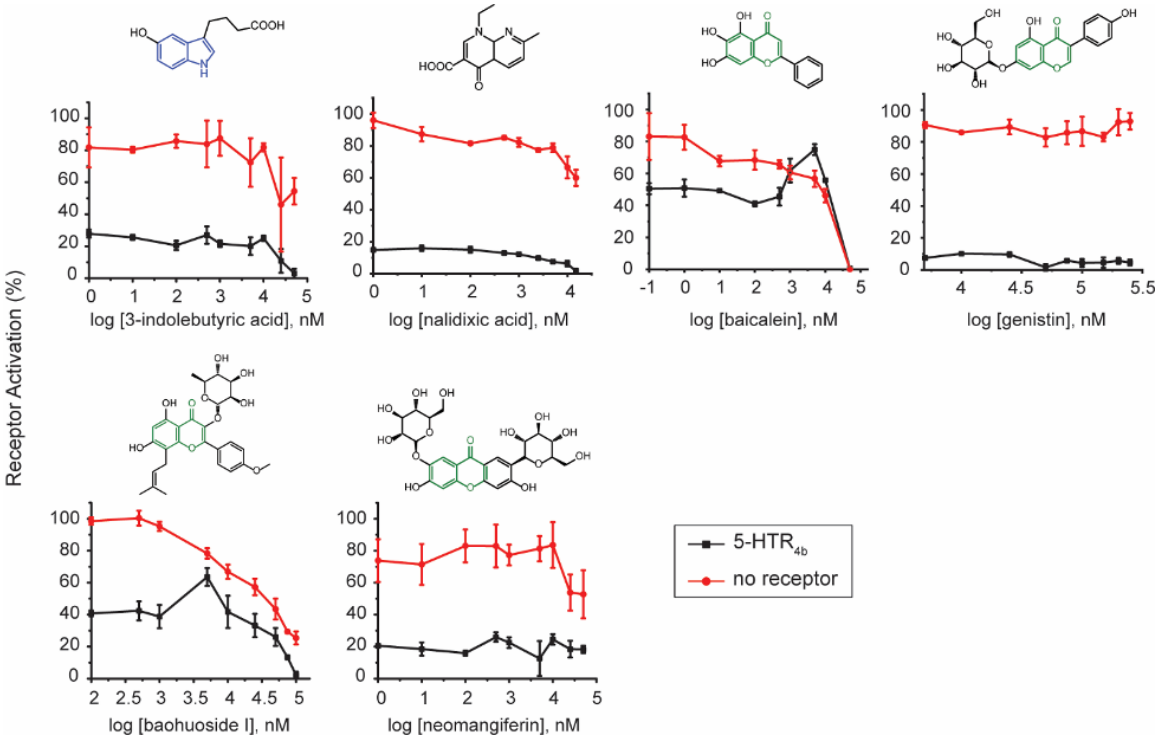

Figure 3.5 – Dose-response curves of the six natural product hits that could not be validated

3.3.5 High-Throughput Screening of a 403-Member Anti-infection Library.

Hordenine has antibiofilm activity against *Pseudomonas aeruginosa*,[30] and halofuginone is used for the treatment and prophylaxis of cryptosporidiosis in ruminants.[31] Intrigued by the fact that antimicrobial agents activate 5-HTR$_{4b}$, we explored whether other antimicrobials activate 5-HTR$_{4b}$. To do this, we screened a commercial 403-member anti-infection library at two chemical concentrations: (i) 10 μM, the same concentration used to screen the natural products library, and (ii) 1 μM due to potential toxicity effects of the anti-infection compounds on yeast. On the basis of *Z*-scores, the 10 μM screen resulted in

four hits: proflavine, ethidium bromide, ethacridine, and hordenine, while the 1 μM screen also resulted in four hits: isepamicin, osalmid, revaprazan, and borneol (Figure 3.6A). We attribute the absence of 1 μM hits in the 10 μM screen to toxicity issues, given that the 5-HTR_{4b} assay is in cells. Neither proflavine, ethacridine, nor revaprazan have been previously shown to activate any serotonin receptor.[29]

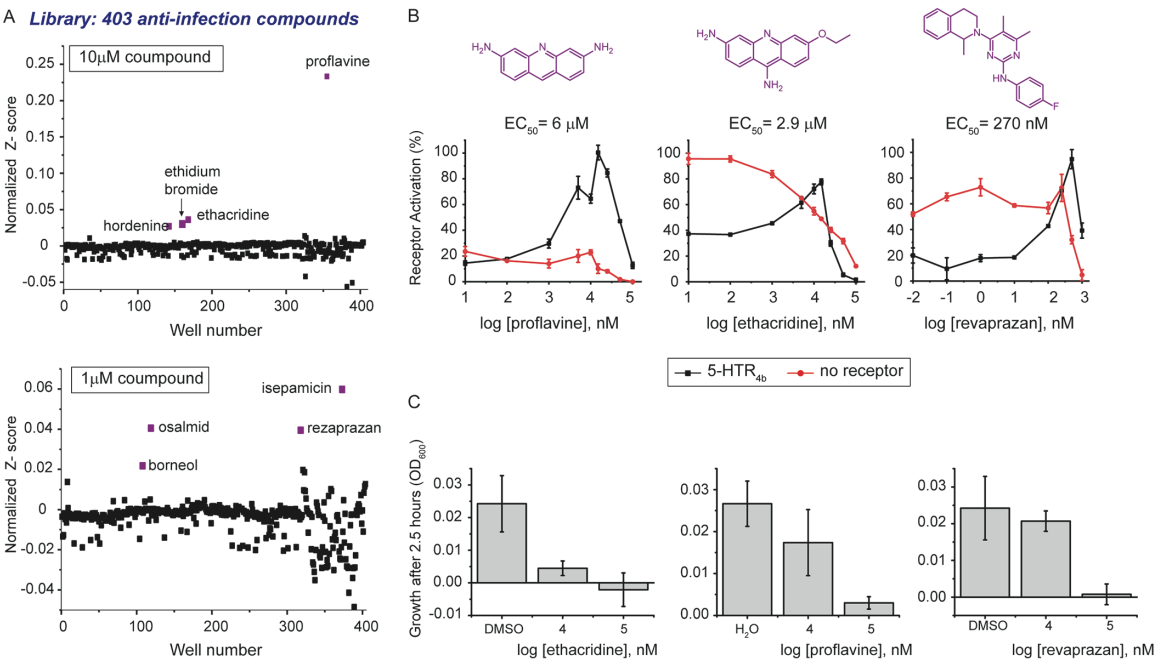

Figure 3.6 – Screening of a 403 member anti-infective library. (A) Anti-infective compound screening results at 10 μM (top) and 1 μM (bottom). The compounds were screened using the 5-HTR_{4b} assay in singlets. *Z*-scores were normalized to the serotonin positive control, which was set to 1 (dotted line). (B) Dose-response curves of the 3 validated anti-infective hits. (C) Toxicity results after yeast was incubated for 2.5 h with varying concentration of the anti-infective hits. Data was collected in triplicate. Shown are means ± s.d.

To confirm the anti-infection library hits, we ran dose-response curves using the 5-HTR_{4b} assay strain, and a control strain carrying the luciferase reporter and a blank plasmid in lieu of 5-HTR_{4b}. We confirmed proflavine (EC_{50} 6 μM), ethacridine (EC_{50} 2.9 μM), and revaprazan (EC_{50} 270 nM) as HTR_{4b} ligands (Figure 3.6B). We noticed a sharp decline in

receptor activation at high compound concentrations. A toxicity assessment of these compounds shows they are toxic at high concentrations (Figure 3.6C). We could not confirm ethidium bromide, isepamicin, osalmid, or borneol as 5-HTR$_{4b}$ ligands (Figure 3.7).

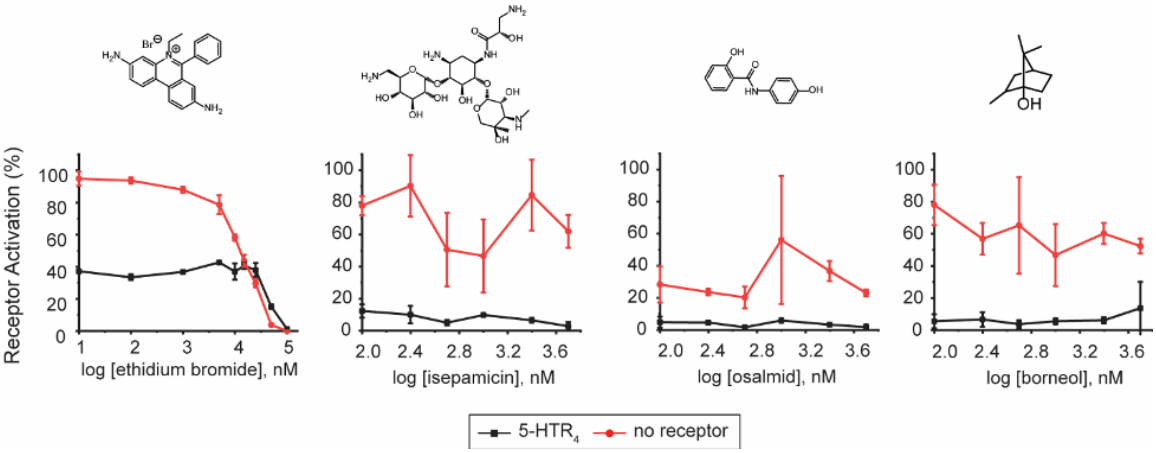

Figure 3.7 – Dose-response curves of the four anti-infection hits that could not be validated

3.3.6 Validating 5-HTR$_{4b}$ Ligands via Wound Healing and Motility Assays in Colon Epithelial Cells.

To assess the biological relevance of the five 5-HTR$_{4b}$ identified ligands, we tested their ability to activate 5-HTR$_{4b}$ in mammalian cells by inducing cell motility and wound healing in human colon epithelial cells (Caco-2). Caco-2 cells endogenously express 5-HTR$_4$[32] and activation of 5-HTR$_4$ using the agonist tegaserod leads to increased cell motility, wound healing, and cell proliferation.[32] Tegaserod (Zelmac) is used for the treatment of IBS-C. To test cell motility, we used culture inserts to create a cell free zone to avoid cell death and damage.[33] Revaprazan and hordenine significantly ($P < 0.01$)

increase cell motility (Figure 3.8A,B). Wound healing was tested by creating a scratch in the cell monolayer using a pipet tip. This more closely mimics *in vivo* wound healing by

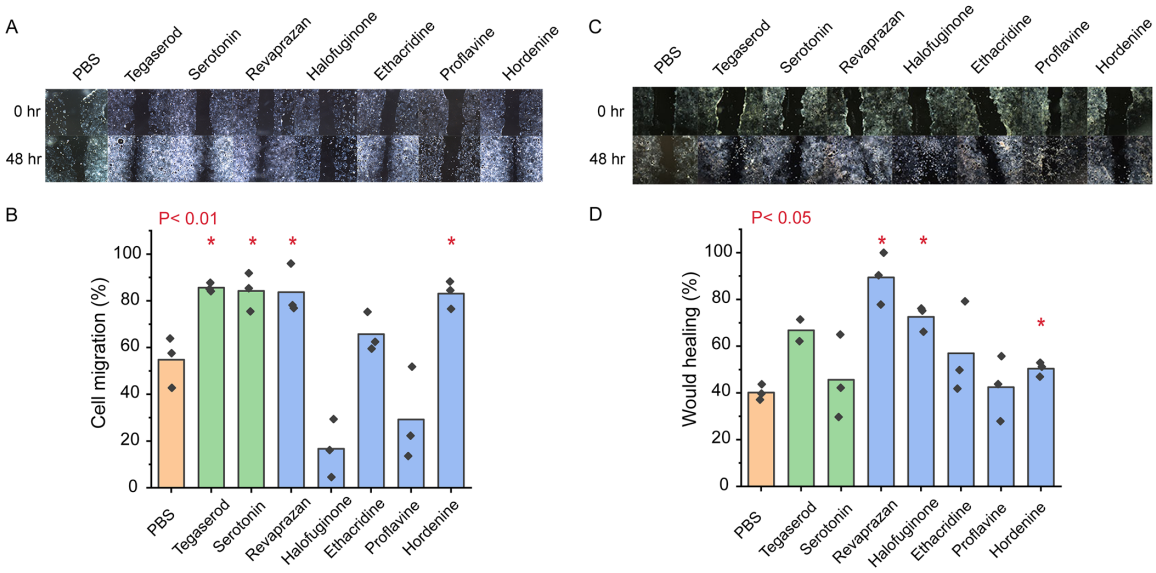

Figure 3.8 – Antimicrobials enhanced colon epithelial cell motility and wound healing. (A) Representative photomicrographs of the colon cell (Caco-2) migration assay after 48 h with phosphate buffer control (PBS), tegaserod (5-HTR$_{4B}$ agonist, 1 μM), serotonin (1 μM), and the five 5-HTR$_{4B}$ ligands (10 μM). For all photomicrographs for migration after 24 and 48 h, see Figure 3.9. (B) Quantification of the colon cell migration assay after 48 h. The cell migration elicited by revaprazan and hordenine is similar to that observed by serotonin and tegaserod, and it is statistically significantly different than the PBS control ($P < 0.01$). (C) Representative photomicrographs of the Caco-2 wound healing scratch assay after 48 h with control (PBS), tegaserod (1 μM), serotonin (10 μM), and the 5 identified 5-HTR$_{4B}$ ligands (10 μM). For all photomicrographs of the wound healing scratch assay after 24 and 48 h, see Figure 3.10. (D) Quantification of the colon cell wound healing scratch assay. The wound healing elicited by revaprazan, halofuginone, and hordenine is statistically significantly different than the PBS control ($P < 0.05$).

creating cell damage, such as increasing reactive oxygen species at the wound boundary.[34] Incubation of Caco-2 cells with hordenine, halofuginone, and revaprazan resulted in a statistically significant increase in wound healing ($P < 0.05$) when compared to the buffer control (Figure 3.8C,D). Proflavine and ethacridine did not result in a significant increase

in either colon cell motility or wound healing. Taken together, revaprazan and hordenine results in both increased cell motility and wound healing, while halofuginone results only in increased wound healing.

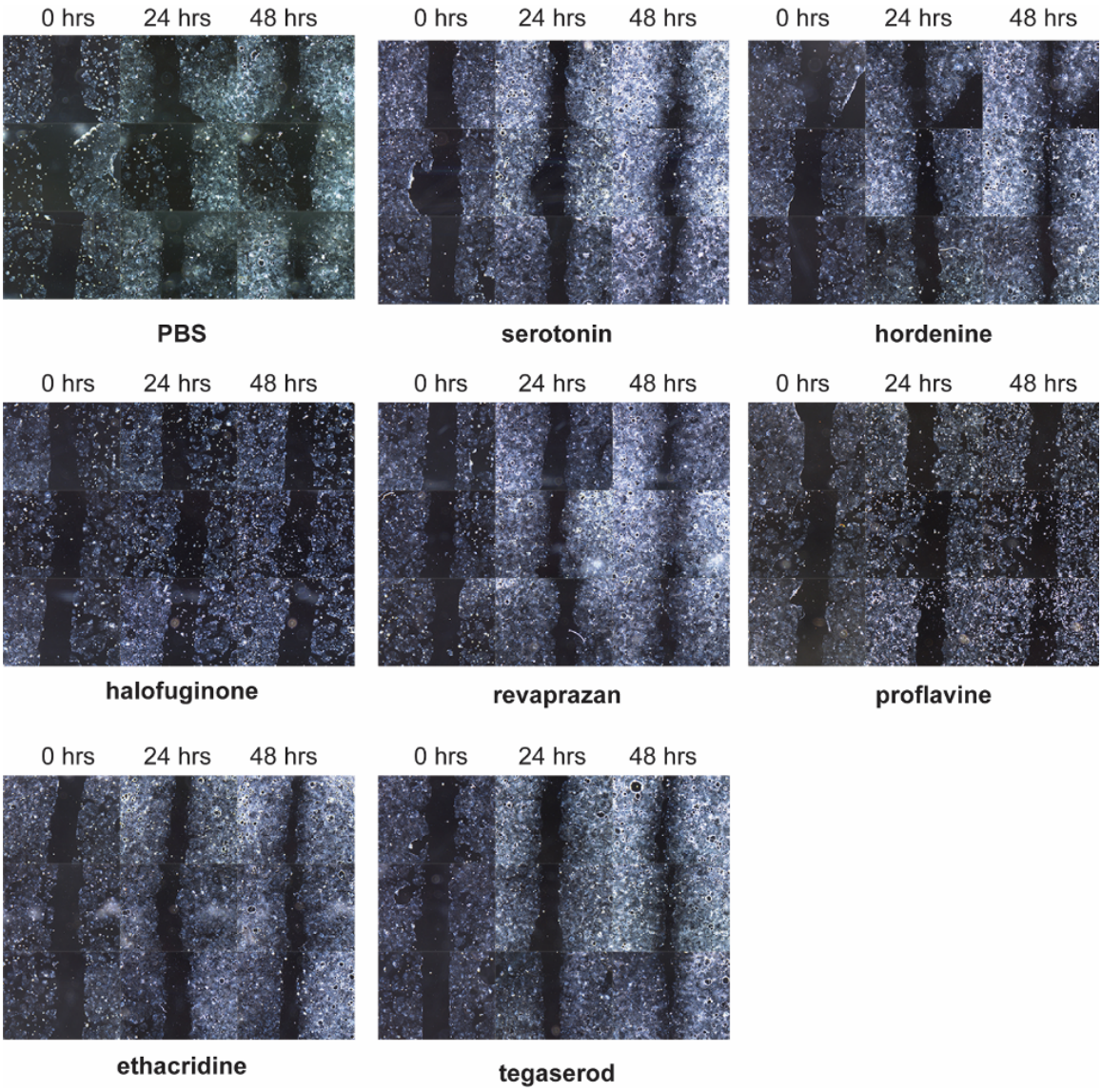

Figure 3.9 – Photomicrographs of colon cell migration assay after 0, 24, and 48 hours

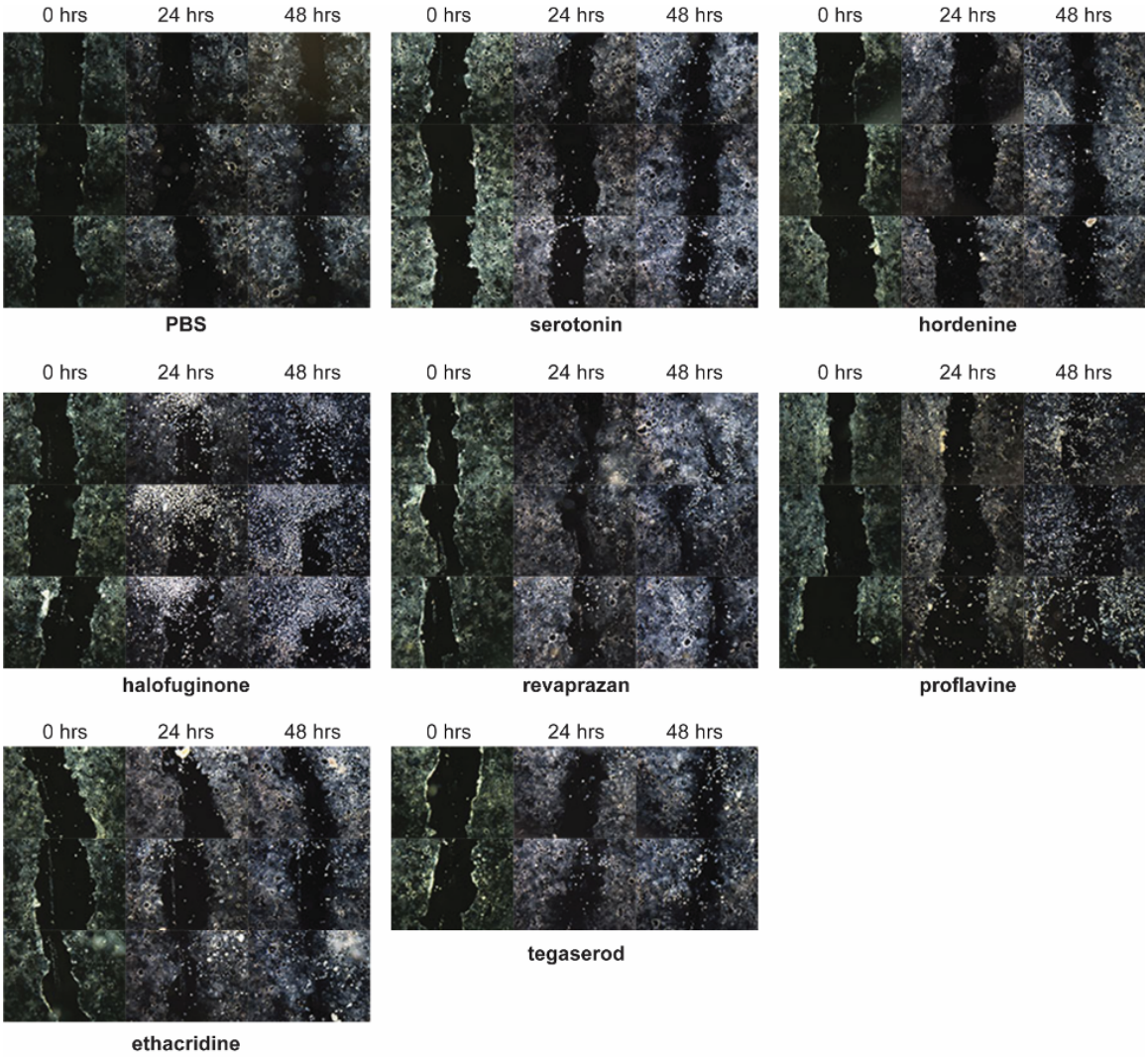

Figure 3.10 – Photomicrographs of colon cell scratch would healing assay after 0, 24, and 48 hours.

3.3.7 Relevance of the Newly Identified 5-HTR$_{4b}$ Agonists in the Gut.

Hordenine is present in malted barley and beer (strong beer: 5.16 mg/L; Pilsner: 2.7 mg/L[35,36]). Assuming all hordenine consumed reaches the colon, a pint of Pilsner beer would need to be ingested to reach concentrations that were observed to increase colon cell motility and wound healing. Other sources of hordenine are athletic performance and weight loss supplements.[37]

Halofuginone is used for the treatment and prophylaxis of cryptosporidiosis in ruminants[31] with a maximum residue limit of 10−30 μg in bovine muscle, fat, kidney, and liver.[38] Chickens treated with halofuginone can expect to lay eggs with up to 60 μg/kg halofuginone.[39] One may only need to eat half a jumbo egg to reach the levels seen to increase colon cell wound healing, assuming all halofuginone reaches the colon.

Revaprazan is a proton pump inhibitor that reduces gastric acid secretion and reduces inflammation caused by *Helicobacter pylori*.[40] At a daily dose of 200 mg/day and assuming it all reaches the colon, the concentration of revaprazan in the gut would be 983.6 μM, which is more than 3000 times higher than the EC_{50} of 270 nM.

3.4 Discussion

We have developed a rapid and robust luciferase-based 5-HTR_{4b} assay in yeast with a 38-fold dynamic range and a limit of detection in the nM level for serotonin. We applied this assay to screen more than 1000 chemicals and identified five new 5-HTR_{4b} ligands. Three of these ligands, hordenine, halofuginone, and revaprazan, activate 5-HTR_{4b} in mammalian cells, eliciting colon cell motility and/or wound healing. The 5-HTR_{4b} assay can now be used to screen large pharmaceutical libraries to identify 5-HTR_{4b} agonists for the treatment of IBS-C. The increased sensitivity of the 5-HTR_{4b} luciferase assay over our previous GFP-based assay should enable the generation of assays for other 5-HTRs, including 5-HTR_1, the pharmaceutical target of antidepressant drugs.[41]

A current limitation of the luciferase-based 5-HTR_{4b} assay is the relative high background in the no receptor control. In the absence of receptor, there is free floating $G_{\alpha\beta\lambda}$, which does not need to be dissociated to activate the kinase cascade leading to increased gene expression.[42] Sequestering $G_{\alpha\beta\lambda}$ by expression of the yeast endogenous GPCR (Ste2) may reduce basal luciferase levels in the control strain. Additionally, mixtures of heterologous GPCR/Ste2 in the sensor strains could achieve lower basal expression, but the system would have to be tuned.

The identification of antimicrobials binding 5-HTR_4 is of profound impact to understanding conditions mediated by serotonin receptors. As shown in this work, antimicrobials may not only interact with the gut microbiota, but also with the human host leading to potential changes in gut movement and secretion. Changes in bowel movements and gastrointestinal hormone secretion may at least be partially caused by activation of serotonin receptors. Thus, our findings have implications for the treatment of IBS-C, which can be treated using antibiotics,[43] and more generally to conditions related to serotonin receptor activation, such as depression. Finally, although antimicrobials have not been extensively considered as activating GPCRs, the availability of this high-throughput GPCR assay will enable the analysis of other GPCRs for their binding of antimicrobials.

3.5 Methods

3.5.1 Materials

Serotonin hydrochloride (S0370) was purchased from TCI Chemicals. RS67333 hydrochloride (SML1882), tegaserod malate (SML1504), metoclopramide

monohydrochloride (M0763), prucalopride (SML1371), cisapride (CDS021610), mosapride citrate (M2946) and zacopride hydrochloride hydrate (SML0081) were purchased from Sigma-Aldrich. Nano-Glo® Luciferase Assay System (N1120) was purchased from Promega. The natural products library (L1400) and the anti-infection library (L3100) were purchased from Selleck Chemicals. Breathe Easy Sealing Membrane was purchased from Electron microscope services. Luminescence white 96-well plates were purchased from Thermo

3.5.2 Luciferase-Based 5-HTR$_{4b}$ Assay Construction.

NanoLuc was codon-optimized for *Saccharomyces cerevisiae*, commercially synthesized and cloned into pKM586[44] between *NcoI* and *NheI* to generate pRS415-Leu2-pFIG1-NanoLuc (pEY15). pEY15 was sequenced verified using primer EY248. To construct the 5-HTR$_{4b}$ assay, pEY15, and pESC-His3-pTEF1-HTR$_4$ (pTMC18[7]) were co-transformed into the PPY140[44] (W303 *Δ*far1 *Δ*ste2, *Δ*sst2) to generate PPY1808. To construct the controls strain lacking 5-HTR$_{4b}$, PPY140 was co-transformed with pEY15 and an empty vector (PPY111) to generate PPY1809.

3.5.3 Luciferase-Based 5-HTR$_{4b}$ Assay.

3.5.3.1 Serotonin Detection.

An overnight culture of PPY1808 was used to inoculate 20 mL of synthetic complete medium with 2% glucose lacking histidine and leucine (SD(HL$^-$) to an OD_{600} = 0.06. After 18 h at 15 °C (150 rpm), PPY1808 was centrifuged (3500 rpm, 10 min), and

resuspended to an OD_{600} = 1. In a white flat bottom 96-well plate, 190 μL of pH = 7 SD(HL^-), 8 μL of PPY1808, and 2 μL of serotonin (final concentrations 0−4.3 mM) were added. After the chemical incubation step (2 h, plates covered with Breathe Easy Sealing Membrane, 30 °C, 250 rpm), 20 μL of 1:100 mixture of NanoLuc substrate to NanoLuc buffer was added[45] for the luciferase substrate incubation step well (30 min, plates covered with Breathe Easy Sealing Membrane, 30 °C, 250 rpm). Luminescence was read immediately after luciferase substrate incubation in a Biotek Synergy 2 plate reader using default settings. For the time course assay, the cells were incubated at medium shaking speed in a Biotek Synergy 2.

3.5.3.2 Screening of Known 5-HTR_4 Agonists.

The same protocol for serotonin detection was followed. Instead of 2 μL of serotonin the same volume of RS67333, zacopride, and metoclopramide in water or cisapride, mosapride, tegaserod, and prucalopride in DMSO were used to reach a final concentrations between 0 and 10000 nM. Dose-response curves were fitted using the dose-response equation in Origin Pro 2016. EC_{50}s were derived from this function by Origin.

3.5.4 High-Throughput Chemical Compound Screening.

3.5.4.1 High-Throughput Assay Validation.

A three-day assay validation was performed according to Iversen*, et al.* [20] The same serotonin detection protocol was followed. High signal wells have 5 μM of serotonin, mid signal wells have 60 nM of serotonin, and low signal wells have 0 nM of serotonin. Data

were analyzed in the Excel template provided by Iversen, *et al.*[20] using activation assay settings with statistics calculated for single replicate screens.

3.5.4.2 Chemical Compound Screening.

The same protocol used for serotonin detection was followed. Instead of 2 μL of serotonin, and 2 μL of chemical in DMSO or water from the natural product library (Selleck Chemicals L1400, 803 chemicals) or the anti-infection compound library (Selleck Chemicals L3100, 403 chemicals) was used to a final concentration of 10 μM (and 1 μM in the case of the anti-infection library). Each plate contained three negative control wells with DMSO or water and three positive controls with serotonin (5 μM) in DMSO. Natural product hits were identified using a *Z*-score as calculated below. The *Z*-scores were normalized to the serotonin positive control in each 96-well plates.

$$z - score = \frac{natural\ product - mean\ DMSO}{STDev\ DMSO}$$

3.5.5 Principal Component Analysis.

3.5.5.1 Data Gathering.

Simplified molecular-input line-entry system (SMILES) for 801 of the 803 chemicals in the natural product library (lanolin and tea polyphenols are mixtures) were retrieved from Selleck Chemicals and PubChem to obtain data on 23 chemical descriptors with 8 descriptors representing chemical functional groups, and 15 descriptors according to Wenderski, *et al.*[46] The SMILES were inputted into Instant JChem (ChemAxon).

3.5.5.2 Data Analysis.

PCA was run using Solo software (Eigenvector Research). Evaluation of the three PCs was conducted using Solo based on Eigenvalues. The three PC scores data from Solo were used to create 3-D scatterplots in MATLAB. The scatterplots were divided into eight distinct chemical spaces based on their 3-D coordinates.

3.5.6 Toxicity assessment

An overnight culture of PPY1808 in SD(HL-) was used adjusted to an OD of 1. In a 96 well plate, 190 µL of pH=7 SD(HL-) , 8 µL of PPY1808, and 2 µL of DMSO, ethracridine in DMSO, proflavine in water, or rezaprazan in DMSO (final concentrations 0, 10µM, 100 µM) were added. The plate was incubated in a Tecan Infinite M200 Pro plate reader at 30 °C with shaking and OD600 was measured at 2.5 hours

3.5.7 Colon Cell Wound Healing Assay.

Caco-2 cells (ATCC HTB-37) were grown at 37 °C with 5% CO_2 in Dulbecco's Modified Eagle Medium (DMEM) with high glucose, sodium pyruvate and GlutaMAX or L-glutamine (Gibco) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin. Media was changed every 2−3 days. Once 70−80% confluence was reached in T25 flasks, cells were detached using 0.05% trypsin−EDTA and diluted to a final volume of 15 mL with growth media. Cells were seeded into a 48 well plate and allowed to grow until they reached 90−100% confluence. Wounds were made with a P200 pipet tip and cells were washed with PBS. Cells were incubated in fresh growth media with either PBS, 10 μM of chemical hit or 1 μM of serotonin or tegaserod. Images were taken on a Leica inverted microscope at 0, 24, and 48 h. ImageJ MRI Wound Healing Tool [http://dev.

mri.cnrs.fr/projects/imagej-macros/wiki/Wound_Healing_Tool] was used to measure wound size.

3.5.8 Cell Migration Assay.

Caco-2 cells were maintained as stated above. Ibidi culture inserts were placed in a 24 well plate, and 70 μL of cells ((1−2) × 10^5 cells/mL) were placed in both insert wells. Cells were grown overnight to allow for attachment. Inserts were removed with sterile tweezers. Cells were washed twice with PBS then incubated with either PBS, 10 μM of chemical hit or serotonin, or 1 μM of tegaserod. Images were taken on a Leica inverted microscope at 0, 24, and 48 h. ImageJ MRI Wound Healing Tool was used to measure the cell free zone.

3.6 References

[1] Gershon, M. D., and Tack, J. (2007) The serotonin signaling system: from basic understanding to drug development for functional GI disorders. *Gastroenterology 132,* 397−414.

[2] Hoffman, J. M., Tyler, K., MacEachern, S. J., Balemba, O. B., Johnson, A. C., Brooks, E. M., Zhao, H., Swain, G. M., Moses, P. L., Galligan, J. J., Sharkey, K. A., Greenwood-Van Meerveld, B., and Mawe, G. M. (2012) Activation of colonic mucosal 5-HT(4) receptors accelerates propulsive motility and inhibits visceral hypersensitivity. *Gastroenterology 142,* 844−854.

[3] Tonini, M. (2005) 5-Hydroxytryptamine effects in the gut: the 3, 4, and 7 receptors. *Neurogastroenterol. Motil.* 17, 637−642.

[4] Wohlfarth, C., Schmitteckert, S., Haïtle, J. D., Houghton, L. A., Dweep, H., Fortea, M., Assadi, G., Braun, A., Mederer, T., Pöhner, S., Becker, P. P., Fischer, C., Granzow, M., Mönnikes, H., Mayer, E. A., Sayuk, G., Boeckxstaens, G., Wouters, M. M., Simreń, M., Lindberg, G., Ohlsson, B., Schmidt, P. T., Dlugosz, A., Agreus, L., Andreasson, A., D'Amato, M., Burwinkel, B., Bermejo, J. L., Röth, R., Lasitschka, F., Vicario, M., Metzger, M., Santos, J., Rappold, G. A., Martinez, C., and Niesler, B. (2017) miR-16 and miR-103 impact 5-HT4 receptor signalling and correlate with symptom profile in irritable bowel syndrome. *Sci. Rep. 7,* 14680.

[5] Sikander, A., Rana, S. V., and Prasad, K. K. (2009) Role of serotonin in gastrointestinal motility and irritable bowel syndrome. *Clin. Chim. Acta 403,* 47−55.

[6] Sipes, N. S., Martin, M. T., Kothiya, P., Reif, D. M., Judson, R. S., Richard, A. M., Houck, K. A., Dix, D. J., Kavlock, R. J., and Knudsen, T. B. (2013) Profiling 976 ToxCast Chemicals across 331 Enzymatic and Receptor Signaling Assays. *Chem. Res. Toxicol. 26,* 878−895.

[7] Ehrenworth, A. M., Claiborne, T., and Peralta-Yahya, P. (2017) Medium-Throughput Screen of Microbially Produced Serotonin via a G-Protein-Coupled Receptor-Based Sensor. *Biochemistry 56,* 5471− 5475.

[8] Hall, M. P., Unch, J., Binkowski, B. F., Valley, M. P., Butler, B. L., Wood, M. G., Otto, P., Zimmerman, K., Vidugiris, G., Machleidt, T., Robers, M. B., Benink, H. A., Eggers, C. T., Slater, M. R., Meisenheimer, P. L., Klaubert, D. H., Fan, F., Encell, L. P., and Wood, K. V. (2012) Engineered luciferase reporter from a deep

sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem. Biol. 7,* 1848−1857.

[9] Nam, Y., Min, Y. S., and Sohn, U. D. (2018) Recent advances in pharmacological research on the management of irritable bowel syndrome. *Arch. Pharmacal Res. 41,* 955−966.

[10] Manabe, N., Wong, B. S., and Camilleri, M. (2010) New-generation 5-HT4 receptor agonists: potential for treatment of gastrointestinal motility disorders. *Expert Opin. Invest. Drugs 19,* 765−775.

[11] Champion, M. C. (1997) Prokinetic therapy in gastroesophageal reflux disease. *Can. J. Gastroenterol. 11,* 55B−65B.

[12] Pascual-Brazo, J., Castro, E., Díaz, A., Valdizań, E. M., Pilar-Cuełlar, F., Vidal, R., Treceño, B., and Pazos, A. (2012) Modulation of neuroplasticity pathways and antidepressant-like behavioural responses following the short-term (3 and 7 days) administration of the 5-HT(4) receptor agonist RS67333. *Int. J. Neuropsychopharmacol. 15,* 631−643.

[13] Costall, B., Domeney, A. M., Gerrard, P. A., Kelly, M. E., and Naylor, R. J. (1988) Zacopride - Anxiolytic Profile in Rodent and Primate Models of Anxiety. *J. Pharm. Pharmacol. 40,* 302−305.

[14] Gralla, R. J., Itri, L. M., Pisko, S. E., Squillante, A. E., Kelsen, D. P., Braun, D. W., Jr, Bordin, L. A., Braun, T. J., and Young, C. W. (1981) Antiemetic efficacy of high-dose metoclopramide: randomized trials with placebo and prochlorperazine in patients with chemotherapy-induced nausea and vomiting. *N. Engl. J. Med. 305,* 905−909.

[15] Bender, E., Pindon, A., van Oers, I., Zhang, Y. B., Gommeren, W., Verhasselt, P., Jurzak, M., Leysen, J., and Luyten, W. (2000) Structure of the human serotonin 5-HT4 receptor gene and cloning of a novel 5-HT4 splice variant. *J. Neurochem. 74,* 478−489.

[16] Mialet, J., Berque-Bestel, I., Sicsic, S., Langlois, L., Fischmeister, R., and Lezoualc'h, F. (2000) Pharmacological characterization of the human 5-HT(4(d)) receptor splice variant stably expressed in Chinese hamster ovary cells. *Br. J. Pharmacol. 131,* 827−835.

[17] Gerald, C., Adham, N., Kao, H. T., Olsen, M. A., Laz, T. M., Schechter, L. E., Bard, J. A., Vaysse, P. J., Hartig, P. R., and Branchek, T. A. (1995) The 5-HT4 receptor: molecular cloning and pharmacological characterization of two splice variants. *EMBO J. 14,* 2806−2815.

[18] Blondel, O., Gastineau, M., Dahmoune, Y., Langlois, M., and Fischmeister, R. (1998) Cloning, expression, and pharmacology of four human 5-hydroxytryptamine 4 receptor isoforms produced by alternative splicing in the carboxyl terminus. *J. Neurochem. 70,* 2252− 2261.

[19] Claeysen, S., Sebben, M., Journot, L., Bockaert, J., and Dumuis, A. (1996) Cloning, expression and pharmacology of the mouse 5-HT(4L) receptor. *FEBS Lett. 398,* 19−25.

[20] Iversen, P. W., Benoit, B., Chen, Y.-F., Dere, W., Devanarayan, V., Eastwood, B. J., Farmen, M. W., Iturria, S. J., Montrose, C., Moore, R. A., Weidner, J. R., and Sittampalam, G. S. (2012) In *Assay Guidance Manual* (Sittampalam, G. S., Coussens, N. P., Brimacombe, K., Grossman, A., Arkin, M., Auld, D., Austin, C.,

Baell, J., Bejcek, B., Chung, T. D. Y., Dahlin, J. L., Devanaryan, V., Foley, T. L., Glicksman, M., Hall, M. D., Hass, J. V., Inglese, J., Iversen, P. W., Kahl, S. D., Kales, S. C., Lal-Nag, M., Li, Z., McGee, J., McManus, O., Riss, T., Trask, O. J., Jr., Weidner, J. R., Xia, M., and Xu, X., Eds.) pp 937−968, Eli Lilly and Company and the National Center for Advancing Translational Sciences, Bethesda, MD.

[21] Zhang, J. H., Chung, T. D. Y., and Oldenburg, K. R. (1999) A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screening 4,* 67−73.

[22] Mishra, A., Dobritsa, S. V., Crouch, M. L., Rabenstein, J., Lee, J. X., and Dhakshinamoorthy, S. (2015) Establishment and validation of a 384-well antibacterial assay amenable for high-throughput screening and combination testing. *J. Microbiol. Methods 118,* 173−175.

[23] Samuel, B. S., Shaito, A., Motoike, T., Rey, F. E., Backhed, F., Manchester, J. K., Hammer, R. E., Williams, S. C., Crowley, J., Yanagisawa, M., and Gordon, J. I. (2008) Effects of the gut microbiota on host adiposity are modulated by the short-chain fatty-acid binding G protein-coupled receptor, Gpr41. *Proc. Natl. Acad. Sci. U. S. A. 105,* 16767−16772.

[24] Maslowski, K. M., Vieira, A. T., Ng, A., Kranich, J., Sierro, F., Yu, D., Schilter, H. C., Rolph, M. S., Mackay, F., Artis, D., Xavier, R. J., Teixeira, M. M., and Mackay, C. R. (2009) Regulation of inflammatory responses by gut microbiota and chemoattractant receptor GPR43. *Nature 461,* 1282−1286.

[25] Singh, N., Gurav, A., Sivaprakasam, S., Brady, E., Padia, R., Shi, H., Thangaraju, M., Prasad, P. D., Manicassamy, S., Munn, D. H., Lee, J. R., Offermanns, S., and

Ganapathy, V. (2014) Activation of Gpr109a, receptor for niacin and the commensal metabolite butyrate, suppresses colonic inflammation and carcinogenesis. *Immunity 40,* 128−139.

[26] Bhattarai, Y., Williams, B. B., Battaglioli, E. J., Whitaker, W. R., Till, L., Grover, M., Linden, D. R., Akiba, Y., Kandimalla, K. K., Zachos, N. C., Kaunitz, J. D., Sonnenburg, J. L., Fischbach, M. A., Farrugia, G., and Kashyap, P. C. (2018) Gut Microbiota-Produced Tryptamine Activates an Epithelial G-Protein-Coupled Receptor to Increase Colonic Secretion. *Cell Host Microbe 23,* 775−785.

[27] Wang, L., Martinez, V., Kimura, H., and Tache, Y. (2007) 5-Hydroxytryptophan activates colonic myenteric neurons and propulsive motor function through 5-HT4 receptors in conscious mice. *Am. J. Physiol. Gastrointest. Liver Physiol. 292,* G419−428.

[28] Gafner, S., Bergeron, C., Batcha, L. L., Reich, J., Arnason, J. T., Burdette, J. E., Pezzuto, J. M., and Angerhofer, C. K. (2003) Inhibition of [3H]-LSD binding to 5-HT7 receptors by flavonoids from Scutellaria lateriflora. *J. Nat. Prod. 66,* 535−537.

[29] Gaulton, A., Hersey, A., Nowotka, M., Bento, A. P., Chambers, J., Mendez, D., Mutowo, P., Atkinson, F., Bellis, L.J., Cibrián-Uhalte, E., Davies, M., Dedman, N., Karlsson, A., Magariños, M. P., Overington, J. P., Papadatos, G., Smit, I., and Leach, A. R. (2017) The ChEMBL database in 2017. *Nucleic Acids Res. 45,* D945−D954.

[30] Zhou, J. W., Luo, H. Z., Jiang, H., Jian, T. K., Chen, Z. Q., and Jia, A. Q. (2018) Hordenine: A Novel Quorum Sensing Inhibitor and Antibiofilm Agent against Pseudomonas aeruginosa. *J. Agric. Food Chem. 66,* 1620−1628.

[31] de Graaf, D. C., Vanopdenbosch, E., Ortega-Mora, L. M., Abbassi, H., and Peeters, J. E. (1999) A review of the importance of cryptosporidiosis in farm animals. *Int. J. Parasitol. 29,* 1269−1287.

[32] Spohn, S. N., Bianco, F., Scott, R. B., Keenan, C. M., Linton, A. A., O'Neill, C. H., Bonora, E., Dicay, M., Lavoie, B., Wilcox, R. L., MacNaughton, W. K., De Giorgio, R., Sharkey, K. A., and Mawe, G. M. (2016) Protective Actions of Epithelial 5-Hydroxytryptamine 4 Receptors in Normal and Inflamed Colon. *Gastroenterology 151,* 933−944.

[33] Poujade, M., Grasland-Mongrain, E., Hertzog, A., Jouanneau, J., Chavrier, P., Ladoux, B., Buguin, A., and Silberzan, P. (2007) Collective migration of an epithelial monolayer in response to a model wound. *Proc. Natl. Acad. Sci. U. S. A. 104,* 15988−15993.

[34] Nikolic, D. L., Boettiger, A. N., Bar-Sagi, D., Carbeck, J. D., and Shvartsman, S. Y. (2006) Role of boundary conditions in an experimental model of epithelial wound healing. *Am. J. Physiol. Cell Physiol. 291,* C68−C75.

[35] Brauers, G., Steiner, I., and Daldrup, T. (2013) Quantification of the biogenic phenethylamine alkaloid hordenine by LC-MS/MS in beer. *Toxichem. Krimtech. 80,* 323−326.

[36] Sommer, T., Dlugash, G., Hübner, H., Gmeiner, P., and Pischetsrieder, M. (2019) Monitoring of the dopamine D2 receptor agonists hordenine and N-

methyltyramine during the brewing process and in commercial beer samples. *Food Chem. 276,* 745−753.

[37] Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vaźquez-Fresno, R., Sajed,T., Johnson,D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., Serra-Cayuela, A., Liu, Y., Mandal, R., Neveu, V., Pon, A., Knox, C., Wilson, M., Manach, C., and Scalbert, A. (2018) HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res. 46,* D608−D617.

[38] Hagren, V., Peippo, P., and Lovgren, T. (2005) *Detecting and Controlling Veterinary Drug Residues in Poultry in Food Safety Control in the Poultry Industry* (Mead, G. C., Ed) 1st ed., pp 44−82, Woodhead Publishing.

[39] Alexander, J., Auđunsson, G. A., Benford, D., Cockburn, A., Cravedi, J.-P., Dogliotti, E., Di Domenico, A., Feŕnandez-Cruz, M. L., Fürst, P., Fink-Gremmels, J., Galli, C. L., Grandjean, P., Gzyl, J., Heinemeyer, G., Johansson, N., Mutti, A., Schlatter, J., van Leeuwen, R., Van Peteghem, C., and Verger, P. (2008) Opinion of the Scientific Panel on Contaminants in the Food chain on a request from the European Commission on cross-contamination of non-target feeding-stuffs by halofuginone hydrobromide authorised for use as a feed additive. *EFSA J. 657,* 1−31.

[40] Lee, J. S., Cho, J. Y., Song, H., Kim, E. H., and Hahm, K. B. (2012) Revaprazan, a novel acid pump antagonist, exerts anti-inflammatory action against Helicobacter pylori-induced COX-2 expression by inactivating Akt signaling. *J. Clin. Biochem. Nutr. 51,* 77−83.

[41] Artigas, F. (2013) Serotonin receptors involved in anti-depressant effects. *Pharmacol. Ther. 137,* 119−131.

[42] Klein, S., Reuveni, H., and Levitzki, A. (2000) Signal transduction by a nondissociable heterotrimeric yeast G protein. *Proc. Natl. Acad. Sci. U. S. A. 97,* 3219−3223.

[43] Rodino-Janeiro, B. K., Vicario, M., Alonso-Cotoner, C., Pascua-Garcia, R., and Santos, J. (2018) A Review of Microbiota and Irritable Bowel Syndrome: Future in Therapies. *Adv. Ther. 35,* 289−310.

[44] Mukherjee, K., Bhattacharyya, S., and Peralta-Yahya, P. (2015) GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synth. Biol. 4,* 1261−1269.

[45] Masser, A. E., Kandasamy, G., Kaimal, J. M., and Andreasson, C. (2016) Luciferase NanoLuc as a reporter for gene expression and protein levels in Saccharomyces cerevisiae. *Yeast 33,* 191−200.

[46] Wenderski, T. A., Stratton, C. F., Bauer, R. A., Kopp, F., and Tan, D. S. (2015) Principal Component Analysis as a Tool for Library Design: A Case Study Investigating Natural Products, Brand-Name Drugs, Natural Product-Like Libraries, and Drug-Like Libraries. *Methods Mol. Biol. 1263,* 225−242.

CHAPTER 4. FUTURE OUTLOOK

4.1 Conclusions and Future Outlook

4.1.1 Expanding our knowledge of ectopic olfactory receptors in the colon

In Chapter 2, a rapid yeast-based OR sensor platform deorphanized OR10S1 and OR2T4 and found an alternative ligand for OR2A7. These ectopic ORs expressed in the colon do not have known functions in the colon yet. With known OR-ligand pairs now available, *in vitro* and *in vivo* experiments can be performed to probe downstream signaling effects to determine possible roles in the colon. This can improve our understanding of disease states and can lead to ORs as drug targets. The yeast-based sensors can be used to screen therapeutics for these conditions. Additionally, the colon is rife with microbial metabolites. OR51E1 and OR51E2 are well-studied colon ORs and are known to interact with our gut microbiota [1,2]. OR2A7, OR2T4, and OR10S1, most likely interact with our gut microbiota as well. The yeast-based OR sensors can be used to screen microbial metabolites to find if any interactions exist, giving us insight into their endogenous role in the colon.

4.1.2 Future screening applications of the serotonin receptor 4 yeast-based assay

In Chapter 3, a serotonin receptor yeast-based assay was improved for HTS applications by simply switching the reporter from GFP to luciferase. This significantly improved the dynamic range and throughput of the sensor and allowed to test large chemical compound libraries for drug discovery applications. 5-HTR_4 is highly expressed

throughout the colon and is the pharmaceutical target for irritable bowel syndrome with constipation (IBS-C). To date, no microbial metabolites have been found to interact with the receptor. The yeast-based assay can be used to explore the microbiome connection to gastrointestinal disorder by screening metabolites and possibly linking certain microbes to disease states. 5-HTR_4 agonists are used to treat IBS-C, however, only two drugs are FDA approved. This is because they can have adverse side effects, such as heart attack, depression, and suicidal ideations. The yeast-based assay can be used to screen large diverse libraries to identify new chemotypes with the potential for less off-target effects.

4.1.3 Improving heterologous GPCR coupling in yeast

In Chapters 2 and 3 the endogenous yeast G_α, GPA1, was used to couple human GPCRs to the yeast mating pathway. As in humans, ORs couple to adenylate cyclase via $G_\alpha$olf, non-functional coupling of the OR to the yeast G_α could be a reason why we could not deorphanize some ORs and most of them had a ~2-3-fold increase in signal after activation. Another reason could be that the library tested was simply too small (57 chemicals) compared to the chemical library sizes used for ORs and more generally GPCR deorphanization that is in the hundreds[3] to hundreds of thousands[4]. Previously, GPA1/human G_α chimeras have been used to couple human GPCRs to the yeast mating pathway. This method has variable success[5, 6]. In this method, the last five amino acids on the α5 helix of the G_α are introduced at the end of GPA1. The α5 helix interacts with transmembrane 5 and 6 (TM 5 and TM 6) of activated GPCRs[7, 8]. From structural studies, we know that 21 amino acids in $G_{\alpha s}$ under 4 Å away from β2AR form approximately 35 interactions[7]. The last five amino acids only account for nine interactions (between 3.0 - 3.8 Å) and could show the lack of robust success in practice. In addition to interactions

between TM 5 and TM 6 of a GPCR and G_α, the intracellular loop 2 also interacts with residues in β1 and β3 strands of the G_α[7]. Engineering the yeast G_α by swapping the whole α5 helix and changing the β1 and β3 strand residues to match human G_αs is a promising method to increase coupling of GPCRs whose G_α is known.

4.2 References

[1] Pluznick, J. L., Protzko, R. J., Gevorgyan, H., Peterlin, Z., Sipos, A., Han, J., Brunet, I., Wan, L. X., Rey, F., Wang, T., Firestein, S. J., Yanagisawa, M., Gordon, J. I., Eichmann, A., Peti-Peterdi, J., and Caplan, M. J. (2013) Olfactory receptor responding to gut microbiota-derived signals plays a role in renin secretion and blood pressure regulation, *Proc. Natl. Acad. Sci. U. S. A. 110*, 4410-4415.

[2] Priori, D., Colombo, M., Clavenzani, P., Jansman, A. J., Lalles, J. P., Trevisi, P., and Bosi, P. (2015) The Olfactory Receptor OR51E1 Is Present along the Gastrointestinal Tract of Pigs, Co-Localizes with Enteroendocrine Cells and Is Modulated by Intestinal Microbiota, *PloS one 10*, e0129501.

[3] Jones, E. M., Jajoo, R., Cancilla, D., Lubock, N. B., Wang, J., Satyadi, M., Cheung, R., de March, C., Bloom, J. S., Matsunami, H., and Kosuri, S. (2019) A Scalable, Multiplexed Assay for Decoding GPCR-Ligand Interactions with RNA Sequencing, *Cell Syst. 8*, 254-+.

[4] Hu, L. Y. A., Tang, P. M., Eslahi, N. K., Zhou, T., Barbosa, J., and Liu, Q. Y. (2009) Identification of Surrogate Agonists and Antagonists for Orphan G-Protein-Coupled Receptor GPR139, *J. Biomol. Screen 14*, 789-797.

[5] Brown, A. J., Dyos, S. L., Whiteway, M. S., White, J. H. M., Watson, M. A. E. A., Marzioch, M., Clare, J. J., Cousens, D. J., Paddon, C., Plumpton, C., Romanos, M. A., and Dowell, S. J. (2000) Functional coupling of mammalian receptors to the yeast mating pathway using novel yeast/mammalian G protein alpha-subunit chimeras, *Yeast 16*, 11-22.

[6] Dong, S. Y., Rogan, S. C., and Roth, B. L. (2010) Directed molecular evolution of DREADDs: a generic approach to creating next-generation RASSLs, *Nat. Protoc. 5*, 561-573.

[7] Rasmussen, S. G. F., DeVree, B. T., Zou, Y. Z., Kruse, A. C., Chung, K. Y., Kobilka, T. S., Thian, F. S., Chae, P. S., Pardon, E., Calinski, D., Mathiesen, J. M., Shah, S. T. A., Lyons, J. A., Caffrey, M., Gellman, S. H., Steyaert, J., Skiniotis, G., Weis, W. I., Sunahara, R. K., and Kobilka, B. K. (2011) Crystal structure of the beta(2) adrenergic receptor-Gs protein complex, *Nature 477*, 549-U311.

[8] Maeda, S., Qu, Q. H., Robertson, M. J., Skiniotis, G., and Kobilka, B. K. (2019) Structures of the M1 and M2 muscarinic acetylcholine receptor/G-protein complexes, *Science 364*, 552-+.

APPENDIX A. SCREENING FOR SEROTONIN RECEPTOR 4 AGONISTS USING GPCR-BASED SENSOR IN YEAST

Yasi, E. A., and Peralta-Yahya, P. *Methods Mol. Biol.* Accepted.

A1.1 Summary

More than 30% of all pharmaceuticals target G-protein coupled receptors (GPCRs). Here we present a GRPC-based screen in yeast to identify ligands for human serotonin receptor 4 (5-HTR_4). Serotonin receptor 4 agonists are used for the treatment of irritable bowel syndrome with constipation. Specifically, the HTR_4-based screen couples activation of 5-HTR_4 on the yeast cell surface to luciferase reporter expression. The HTR_4-based screen has a throughput of one compound per second allowing the screening of more than a thousand compounds per day.

A1.2 Introduction

There are over 800 G-protein coupled receptors (GPCRs) expressed in humans.[1] These receptors have a plethora of functions from sensing odorants[2] and flavors[3] to binding neurotransmitters[4] to controlling blood pressure and heart rate[5]. Currently, approximately 34% of FDA approved drugs target GPCRs.[6] Notably, levodopa targets dopamine receptors in the brain to treat Parkinson's disease,[7] sumatriptan targets serotonin receptors 1B and 1D to treat migraines,[8] and propranolol targets beta adrenergic receptors to treat hypertension.[9] Some over-the-counter supplements also target GPCRs, such as melatonin, a sleep aid.[10] With the vast consequences that GPCRs have on human health and

physiology, high-throughput assays are needed to rapidly identify ligands, synthetic and endogenous, to further understand their downstream effects.

Mammalian-based GPCR assays have long been used to detect GPCR activation. Indeed, several commercial kits are available to read out endogenous GPCR activation, including cAMP accumulation, which can be linked to cell fluorescence (e.g. CatchPointTM cAMP, Molecular Devices), calcium flux (e.g. FLIPR® Calcium Assay Kits, Molecular Devices), or luminescence.[11-14] A general challenge of using mammalian cell-based assays for the identification of GPCR ligands is that mammalian cells endogenously express tens of GPCRs (e.g. 75 different GPCRs in HEK293).[15] Thus, an assay phenotype can be trigged by a ligand that activates a different GPCR that transduces via the same G*α*. Additionally, mammalian-based assays can be easily contaminated and are inherently slow due to their long doubling time and need for passaging.

The yeast *Saccharomyces cerevisiae* is an ideal host for the development of GPCR-based assays. Haploid yeasts have only two GPCRs: Ste2 (Mat*α*) or Ste3 (Mat*a*) respond to pheromones resulting in yeast mating, and Gpr1 responds to glucose.[16] Detection of GPCR activation can be achieved by expressing a heterologous GPCR on the yeast cell surface, coupling it to the yeast mating pathway and expressing a reporter gene (e.g. green fluorescent protein, luciferase) under control of a mating pathway responsive promoter. To increase the heterologous GPCR signal three deletions are needed. Deletion of *ste2/3*[17] increases signal transduction, and deletions of *far1* and *sst2* prevent cell cycle arrest and desensitization of the Gα subunit, respectively.[18] The receptor Gpr1 couples to a distinct adenylate cyclase cascade that will prevent cross talk with the mating pathway.[19] Traditionally, the reporter gene has been an auxotrophic marker that enables growth

selection,[20,21] which takes up to three days to see results of the assay. More recently, reporters that allow assay read outs in the same day, including green fluorescent protein (4 hours),[22,23] and luciferase (1-2.5 hours),[24,25] have been developed.

Recently, we presented a GPCR-based assay that links serotonin receptor 4b (5-HTR_{4b}) to luciferase (NanoLuc)[26] expression[24]. 5-HTR_{4b} is expressed in the brain but is more prominently expressed throughout the gastrointestinal tract,[27] where over 95% of serotonin is found.[28] 5-HTR_4 is a drug target for irritable bowel syndrome with constipation and currently only a few FDA approved treatments target 5-HTR_4, such as prucalopride (Motegrity™)[29] and metoclopramide (Reglan®)[30]. The luciferase-based 5-HTR_4 assay is fast, going from chemical incubation to signal read out in 2.5 hours. The ability to use a luminescence plate reader to detect reporter gene expression enables a screening rate of 1 well per second.[24] This method enables the high-throughput screening of vast chemical libraries for 5-HTR_{4b} activation.

A1.3 Materials

i. Plasmids and Strains

1. GPCR plasmid: PPY1192: pESC-His3-P_{TEF1}-HTR_{4b}-T_{Cyc1}
2. No GPCR plasmid: PPY111: pESC-His3-P_{TEF1}-T_{Cyc1}
3. GPCR microscopy plasmid: PPY1133: pESC-His3-P_{TEF1}-HTR_{4b}-GFP-T_{Cyc1}
4. Reporter Plasmid: PPY1740: pRS415-Leu2-P_{FIG1}-NanoLuc
5. *Yeast sensor strain*: PPY140: *S. cerevisiae* W303-1a MATa *ade2-1*, *ura3-1*, *his3-11*, *trp1-1*, *leu2-3*, *leu2-112*, *can1-100*, *Δfar1*, *Δsst2*, *Δste2*
6. *5-HTR_{4b}-sensing strain*: PPY140 carrying plasmids PPY1192 + PPY1740

7. *Control sensing strain*: PPY140 carrying plasmids PPY111 + PPY1740
8. *5-HTR$_{4b}$-GFP fusion strain*: PPY140 carrying plasmid PPY1133

ii. Yeast Media and Transformation Supplies

1. Synthetic Dropout media lacking histidine and leucine with 2% glucose (SD (HL^-))[31]
2. Synthetic Dropout media lacking histidine and leucine with 2% glucose[31] and buffered with potassium phosphate to pH 7 (SD(HL^-) pH 7)
3. Synthetic Dropout media lacking histidine with 2% glucose[31] (SD(H^-))
4. Yeast Peptone Dextrose media (YPD)[32]
5. Tris-EDTA buffer (10x, pH 7.5): 100 mM Tris, 10 mM EDTA
6. 1 M lithium acetate
7. PEG 4000 50% (w/v)
8. *Solution 1*: 8 mL diH_2O, 1 mL 10x TE buffer, 1 mL 1M lithium acetate
9. *Solution 2*: 800 μL 50% (w/v) PEG 4000, 100 μL 10x TE buffer, 100 μL 1M lithium acetate
10. 10% KOH in water
11. Salmon Sperm DNA: 1mg/mL in water, filter sterilized
12. 1.5 mL Microcentrifuge tubes
13. Petri dishes (100 x 15 mm)
14. Test tubes (25 x 150 mm)
15. Inoculation loops
16. Flat toothpicks
17. Parafilm

iii. Fluorescent Microscopy

1. Poly-L-lysine slides with coverslip
2. Calcofluor White Stain
3. Confocal microscope

iv. Luminescent Plate Reader Assay

1. White opaque flat bottom 96 well plate
2. Chemical library
3. Chemical carrier solvent: e.g. Dimethyl Sulfoxide (DMSO).
4. Breathe-Easy® membrane (Electron Microscopy Sciences)
5. Promega Nano-Glo® Assay System (N1120)
6. Luminescent enabled plate reader

v. Software

1. Microsoft Excel or similar program

A1.4 Methods

i. Preparation of *5-HTR$_{4B}$-sensing, Control sensing,* and *5-HTR$_{4B}$-GFP fusion* strains

1. Streak out PPY140 on a YPD agar plate and incubate at 30°C for 3 days.
2. Patch a single colony on a fresh YPD agar plate using a sterile flat toothpick or a sterile loop and incubate at 30°C for 3 days.

3. Inoculate a small loopful of PPY140 into 5 mL YPD in a test tube. Grow overnight at 30 °C at 250 rpm.
4. Add 1 mL of PPY140 overnight culture to 49 mL of fresh YPD and grow at 30 °C, 250 rpm to an OD_{600} = ~0.6-0.8 which takes approximately 3-4 h.
5. Centrifuge culture 3 min x 1800 g, decant, and resuspend cells in 50 mL diH_2O.
6. Centrifuge again 3 min x 1800 g, decant, and resuspend cells in 5 mL *Solution 1*.
7. Centrifuge again 3 min x 1800 g, decant, and resuspend in 300 μL *Solution 1*.
8. To generate the HTR_{4b} sensing strain and empty vector control, combine in a microcentrifuge tube:
    - 300 μL of *Solution 2*
    - 50 μL of cells in *Solution 1* (*Step 7*)
    - 5 μL of salmon sperm DNA (denatured at 95 °C for 5 mins, chilled on ice afterwards)
    - 500 ng of the GPCR plasmid (replace with No GPCR plasmid for control)
    - 500 ng of the Reporter Plasmids
9. Incubate at 30°C, 250 rpm for 30 minutes. Heat shock the yeast by placing the tubes in a 42°C for 15 min. Spin cells down 1 min x 4,000 RPM.
10. Resuspend the cell pellet in 150 μL of diH_2O and plate on an SD(HL^-) agar plate. Incubate plate at 30°C for 3 days [see **Notes 1, 2**].
11. To generate the *HTR_{4b}-GFP fusion strain* to test GPCR expression in yeast using fluorescent microscopy: use the same protocol as in *Step 8* except replace GPCR plasmid with GPCR microscopy plasmid and do not use Reporter Plasmid. Plate on a SD(H^-) agar plate. Incubate plate at 30°C for 3 days [see **Notes 1, 2**].

12. Patch cells from a single colony on a fresh SD(H^-) agar plate using a sterile flat toothpick or a sterile loop. Incubate plates at 30°C for 3 days.
13. Patches can be stored at 4°C for up to one month wrapped in parafilm.

ii. Fluorescent Microscopy

1. Start an overnight culture of the *HTR_{4b}-GFP fusion strain* by taking a small loopful of the patch and inoculating 5 mL of SD(H^-). Shake at 30°C and 250 rpm. [see **Note 3**].
2. The next day, use the overnight media to inoculate 20 mL of SD(H^-) to an OD_{600} of 0.06. Incubate the culture at 15 °C, 150 rpm for 18h.
3. Centrifuge culture 10 min x 3500 rpm, decant, and resuspend cells in 200 μL SD(H^-).
4. To a microscope slide, layer:
    - 2 μL of the cells collected in *Step 3*
    - 2 μL of calcofluor white stain
    - 2 μL of potassium hydroxide solution.

    Add a cover slide, making sure to avoid air bubbles.
5. Visualize the yeast on a confocal microscope using a 63x objective lens. The microscope should be equipped with both a 488 and a 405 nm laser lines to excited GFP and calcofluor white, respectively.

iii. GPCR Luminescence Assay

1. Start an overnight culture of *5-HTR_{4b}-sensing strain* by taking a small loopful of the patch and inoculating 5 mL of SD(HL^-). Shake at 30°C and 250 rpm. [see **Note 3**].

2. The next day, use the overnight to inoculate 20 mL of SD(H^-) to an OD_{600} = 0.06. Incubate the culture at 15°C, 150 rpm for 18 h.
3. Centrifuge the culture 10 min x 3500 rpm, decant, and resuspend cells in SD(HL^-) to OD_{600} = 1.
4. To a 96 well plate, add:
    - 190 μL of fresh SD(HL^-) pH 7 [see **Note 4**]
    - 8 μL of cell suspension obtained from *Step 3*
    - 2 μL of chemical or chemical carrier solvent as a control [see **Note 5, 6**]. A chemical carrier solvent control and a serotonin positive control should be run in triplicate on the same plate.
    - Cover plate using breathe-easy® membrane [see **Note 7**].
    - Incubate for 2 h x 250 rpm at 30 °C
5. While incubating, allow Promega Nano-Glo® buffer to melt at room temperature. Just before the 2 h incubation is over, create a 1:100 dilution of Nano-Glo® *Substrate*: Nano-Glo® *Buffer*.
6. Remove breathe-easy® membrane and pipet 20 μL of Nano-Glo® *Substrate*: Nano-Glo® *Buffer* dilution into each well.
7. Place a new breathe-easy® membrane and incubate at 30 °C, 250 rpm for 30 min.
8. Remove breathe-easy® membrane and read immediately on luminescent plate reader (i.e. Biotek Synergy 2) [see **Note 8**]

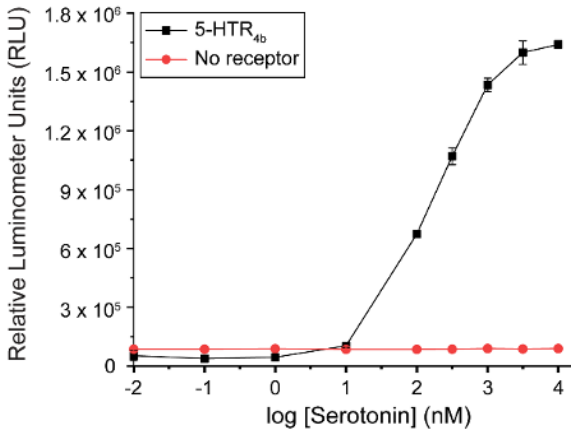

Figure A1.1 – Representative raw data from the 5-HTR$_4$ assay with serotonin.

iv. Secondary Assay: Dose-response Curves

If hits are detected in a screen, follow the same protocol using both the 5-HTR4B-sensing and Control sensing strain and run dose-response curves to confirm luminescence signal is GPCR dependent [see **Note 9**]. Figure A1.1 shows representative data.

v. *Z*-score Analysis

1. Use the equation below to calculate *Z*-scores from the screen:

$$\boldsymbol{Z - score = \frac{(chemical - mean\ vehicle\ control)}{STDev\ vehicle\ control}}$$

2. Normalize the *Z*-scores to the serotonin positive control by dividing the chemical *Z*-score by serotonin's *Z*-score.

A1.5 Notes

1. Amount of yeast transformation mixture that is plated may need to be varied to see single, well-spaced colonies.

2. Colonies should be pink and approximately 2-5 mm in diameter. Smaller white colonies are indicative of satellite colonies that will not grow well when patched or inoculated overnight.
3. Overnight OD_{600}s should be ~6-8.
4. Media is pH to 7 as that is the optimal pH for NanoLuc activity.[26]
5. Chemical carrier solvent should be chosen based on chemical solubility. When using DMSO, use at less ≤ 1% final concentration. If using other carrier solvents, measure the toxicity to yeast to determine maximal final concentration.
6. The 5-HTR_{4b} assay chemical screen can be run in singlets using all wells in the plate.
7. After placement of breathe-easy® membrane, care should be taken to avoid rough handling to prevent cross contamination of wells. Shaking at 250 rpm does not cause well cross contamination.
8. On a Biotek Synergy 2, we use default luminescence gain settings. Testing different gains should be done on each machine to ensure optimal signal read out and avoid signal saturation.
9. The *Control Sensing strain* has higher background than the *5-HTR_{4b}-sensing strain* at low chemical concentrations. The signal of the *5-HTR_{4b}-sensing strain* should be higher than the *Control Sensing strain* with increase in chemical concentration.

A1.6 References

[1] Lv, X. C., Liu, J., Shi, Q. et al. (2016) In vitro expression and analysis of the 826 human G protein-coupled receptors. *Protein Cell* **7**: 325-337.

[2] Dryer, L., Berghard, A. (1999) Odorant receptors: a plethora of G-protein-coupled receptors. *Trends Pharmacol Sci* **20**: 413-417.

[3] Hoon, M. A., Adler, E., Lindemeier, J. et al. (1999) Putative mammalian taste receptors: A class of taste-specific GPCRs with distinct topographic selectivity. *Cell* **96**: 541-551.

[4] Waxham M, N. (2014) Neurotransmitter receptors. In: Byrne JH, Heidelberger, R., Waxham, M. N. (eds) From molecules to networks, Academic Press, pp 285-321.

[5] Salazar, N. C., Chen, J., Rockman, H. A. (2007) Cardiac GPCRs: GPCR signaling in healthy and failing hearts. *Biochim Biophys Acta, Biomembranes* **1768**: 1006-1018.

[6] Hauser, A. S., Attwood, M. M., Rask-Andersen, M., et al. (2017) Trends in GPCR drug discovery: new agents, targets and indications. *Nat Rev Drug Discov* **16**: 829-842.

[7] Hisahara, S., Shimohama, S. (2011) Dopamine Receptors and Parkinson's Disease. *Int J Med Chem* **2011**: 403039.

[8] Akin, D., Gurdal, H. (2002) Involvement of 5-HT1B and 5-HT1D receptors in sumatriptan mediated vasocontractile response in rabbit common carotid artery. *Br J Pharmacol* **136**: 177-182.

[9] Prichard, B. N. C. (1982) Propranolol and Beta-Adrenergic-Receptor Blocking-Drugs in the Treatment of Hypertension. *Br J Clin Pharmacol* **13**: 51-60.

[10] Arendt, J., Rajaratnam, S. M. W. (2008) Melatonin and its agonists: an update. *Br J Psychiatry* **193**: 267-269.

[11] Zhang, R., Xie, X. (2012) Tools for GPCR drug discovery. *Acta Pharmacol Sin* **33**: 372-384.

[12] Chen, Y., Xu, Z., Wu, D. et al. (2015) Luciferase reporter gene assay on human 5-HT receptor: which response element should be chosen? *Sci Rep* **5**: 8060.

[13] Cheng, Z., Garvin, D., Paguio, A. et al. (2010) Luciferase Reporter Assay System for Deciphering GPCR Pathways. *Curr Chem Genomics* **4**: 84-91.

[14] Kroeze, W. K., Sassano, M. F., Huang, X. P., et al. (2015) PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat Struct Mol Biol* **22**: 362-U328.

[15] Atwood, B. K., Lopez, J., Wager-Miller, J., et al. (2011) Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. *BMC genomics* **12**: 14.

[16] Versele, M., Lemaire, K., Thevelein, J. M. (2001) Sex and sugar in yeast: two distinct GPCR systems. *Embo Rep* **2**: 574-579.

[17] Price, L. A., Strnad, J., Pausch, M. H., Hadcock, J. R. (1996) Pharmacological characterization of the rat A2a adenosine receptor functionally coupled to the yeast pheromone response pathway. *Mol Pharmacol* **50**: 829-837.

[18] Price, L. A., Kajkowski, E. M., Hadcock, J. R., et al. (1995) Functional coupling of a mammalian somatostatin receptor to the yeast pheromone response pathway. *Mol Cell Biol* **15**: 6188-6195.

[19] Harashima, T., Heitman, J. (2002) The G alpha protein Gpa2 control yeast differentiation by ineracting with Kelch repeat proteins that mimic G beta subunits. *Mol Cell* **10**: 163-173.

[20] Pausch, M. H. (1997) G-protein-coupled receptors in Saccharomyces cerevisiae: high-throughput screening assays for drug discovery. *Trends Biotechnol* **15**: 487-494.

[21] Pajot-Augy, E., Crowe, M., Levasseur, G., et al. (2003) Engineered yeasts as reporter systems for odorant detection. *J Recep Sig Transd* **23**: 155-171.

[22] Shaw, W. M., Yamauchi, H., Mead, J. et al. (2019) Engineering a Model Cell for Rational Tuning of GPCR Signaling. *Cell* **177**: 782-796.

[23] Mukherjee, K., Bhattacharyya, S., Peralta-Yahya, P. (2015) GPCR-Based Chemical Biosensors for Medium-Chain Fatty Acids. *ACS Synth Biol* **4**: 1261-1269.

[24] Yasi, E. A., Allen, A. A., Sugianto, W., Peralta-Yahya, P. (2019) Identification of Three Antimicrobials Activating Serotonin Receptor 4 in Colon Cells. *ACS Synth Biol* **8**: 2710-2717.

[25] Laschet, C., Dupuis, N., Hanson, J. (2019) A dynamic and screening-compatible nanoluciferase-based complementation assay enables profiling of individual GPCR-G protein interactions. *J Biol Chem* **294**: 4079-4090.

[26] Hall, M. P., Unch, J., Binkowski, B. F., et al. (2012) Engineered luciferase reporter from a deep sea shrimp utilizing a novel imidazopyrazinone substrate. *ACS Chem Biol* **7**: 1848-1857.

[27] Ray, A. M., Kelsell, R. E., Houp, J. A. et al. (2009) Identification of a novel 5-HT4 receptor splice variant (r5-HT4c1) and preliminary characterisation of specific 5-HT4a and 5-HT4b receptor antibodies. *Eur J Pharmacol* **604**: 1-11.

[28] Gershon, M. D., Tack, J. (2007) The serotonin signaling system: from basic understanding to drug development for functional GI disorders. *Gastroenterology* **132**: 397-414.

[29] Vijayvargiya, P., Camilleri, M. (2019) Use of prucalopride in adults with chronic idiopathic constipation. *Expert Rev Clin Phar* **12**: 579-589.

[30] Avalos, D. J., Sarosiek, I., Loganathan, P., McCallum, R. W. (2018) Diabetic gastroparesis: current challenges and future prospects. *Clin Exp Gastroenter* **11**: 347-363.

[31] Cold Spring Harbor Protocols (2006) Complete minimal (cm) or synthetic complete (sc) and drop-out media. 2006(1):pdb.rec8190.

[32] Cold Spring Harbor Protocols (2010) YPD media. 2010(9):pdb.rec12315-pdb.rec12315.

APPENDIX B. PLASMID TABLES

Table B.1: Chapter 2 Plasmids

Number	Name	Description	Source
PPY111	pKM111	pESC-His3-P_{TEF1}-P_{ADH1}	Mukherjee et al. (2015)
PPY586	pKM586	pRS415- P_{FIG1}-eGFP-Leu2	Mukherjee et al. (2015)
PPY1658	pHW3	pESC-HIS3-P_{TEF1}-OR2A42- P_{ADH1}	This study
PPY1661	pHW6	pESC-HIS3-P_{TEF1}-OR2W3- P_{ADH1}	This study
PPY1663	pHW7	pESC-HIS3-P_{TEF1}-OR2T4- P_{ADH1}	This study
PPY1715	pHW18	pESC-HIS3-P_{TEF1}-OR51B5- P_{ADH1}	This study
PPY1795	pPB8	pESC-HIS3-P_{TEF1}-OR2L13- P_{ADH1}	This study
PPY1717	pHW20	pESC-HIS3-P_{TEF1}-OR10S1- P_{ADH1}	This study
PPY1718	pHW21	pESC-HIS3-P_{TEF1}-OR2A7- P_{ADH1}	This study
PPY1719	pHW22	pESC-HIS3-P_{TEF1}-OR2A42-GFP	This study
PPY1720	pHW23	pESC-HIS3-P_{TEF1}-OR51B5-GFP	This study
PPY1734	pHW30	pESC-HIS3-P_{TEF1}-OR2T4-GFP	This study
PPY1737	pHW33	pESC-HIS3-P_{TEF1}-OR2W3-GFP	This study
PPY1947	pPB59	pESC-HIS3-P_{TEF1}- OR2A7-GFP	This study
PPY1948	pPB60	pESC-HIS3-P_{TEF1}-OR10S1-GFP	This study
PPY1930	pPB54	pESC-HIS3-P_{TEF1}-OR2L13-GFP	This study

Table B.2: Chapter 3 Plasmids

Plasmid Number	Plasmid Name	Description	Citation
PPY111	pKM111	pESC-His3-pTEF1-tCYC1	Mukherjee et al. 2015
PPY586	pKM586	pRS415-Leu2-pFIG1-GFP	Mukherjee et al. 2015
PPY1192	pTMC18	pESC-His3-pTEF1-HTR4	Ehrenworth et al. 2017
PPY1740	pEY15	pRS415-Leu2-pFIG1-NanoLuc	This study

APPENDIX C. PRIMER TABLES

Table C.1: Chapter 2 Primers

Name	Sequence
EY46	ACTTCTTGCTCATTAGAAAG
HW12	CTTTTCGGTTAGAGCGGATC
EY270	CTGTCAGCCGTTACATTATCC
EY271	CCACAATACAGCAGACGAAAG
EY272	CCGTGAAGTTCAGCGTAAG
EY273	TTCAGCATCGGGTTAAACAG
EY274	TTCAGACAGAGCAAGCACC
EY275	ACGAAAATGACGACACACAG
ACT-F	TTCTGAGGTTGCTGCTTTGG
ACT-R	ACCGACGATAGATGGGAAGAC
HW1	GCAATCTAATCTAAGTTTTAATTACAAAGGATCC
HW4	AGAATGAGATTCCTTGCCC
HW7	TTCCTTGCCTAGTTCTCTCTTACC
HW8	TTCCATTGCCTTTTGGAAGGC
HW9	GGTACCAATTCTATGGGTGGTG
PB89	TTCCTTCAGGAAGCTGAAGATACC
PB116	CGGCGGCGGGCTA
PB118	CAGTGCACGTTCAACACCC
HW15	AGGAGGGCTTTGGGCAAGGAATCTCATTCTGTGAGCAAGGGCGAGG
HW18	TTGTTGGGTAAGAGAGAACTAGGCAAGGAAGTGAGCAAGGGCGAGG
HW19	GTTGAACCAGCCTTCCAAAAGGCAATGGAAGTGAGCAAGGGCGAGG
HW20	CATTTGTTCACCACCCATAGAATTGGTACCGTGAGCAAGGGCGAGG
PB88	GTGTTCGGTATCTTCAGCTTCCTGAAGGAAGTGAGCAAGGGCGAG
PB115	AGCACCGCAGGTAGCCCGCCGCCGGTGAGCAAGGGCGAGG
PB117	GTTCTGGGTGTTGAACGTGCACTGGTGAGCAAGGGCGAGG

Table C.2: Chapter 3 Primer

Primer Name	Sequence

EY248	gctgttgaaataacaaagacattgg

APPENDIX D. STRAIN TABLES

Table D.1: Chapter 2 Strains

Strain Number	Description	Source
PPY140	*S. cerevisiae* MATa ade2-1 ura3-1 his3-11 trp1-1 leu2-3 leu2-112 can1-100 Δ*far1,* Δ*sst2, Δste2*	Mukherjee et al. (2015)
PPY1800	PPY140 carrying pKM111 and pKM586	This study
PPY1801	PPY140 carrying pHW3 and pKM586	This study
PPY1802	PPY140 carrying pHW6 and pKM586	This study
PPY1803	PPY140 carrying pHW7 and pKM586	This study
PPY1804	PPY140 carrying pHW18 and pKM586	This study
PPY1805	PPY140 carrying pPB8 and pKM586	This study
PPY1806	PPY140 carrying pHW20 and pKM586	This study
PPY1807	PPY140 carrying pHW21 and pKM586	This study
PPY1949	PPY140 carrying pHW22	This study
PPY1950	PPY140 carrying pHW23	This study
PPY1951	PPY140 carrying pHW30	This study
PPY1952	PPY140 carrying pHW33	This study
PPY1953	PPY140 carrying pPB59	This study
PPY1954	PPY140 carrying pPB60	This study
PPY1955	PPY140 carrying pPB54	This study

Table D.2: Chapter 3 Strains

Strain Number	Description	Citation
PPY140	*S. cerevisiae* W303 MATa ade2-1 ura3-1 his3-11 trp1-1 leu2-3 leu2-112 can1-100 Δ*far1* Δ*ste2* Δ*sst2*	Mukherjee et al. 2015
PPY1808	PPY140 carrying pTMC18 and pEY15	This study
PPY1809	PPY140 carrying pKM111 and pEY15	This study

APPENDIX E. PRINCIPAL COMPONENT ANALYSIS TABLES AND FIGURE

Table E.1: Chapter 2 Abbreviations and chemical descriptors for PCA analysis

Descriptors	Abbreviations
Molecular Weight	MW
Carbon atom count	CAC
Hydrogen atom count	HAC
Oxygen atom count	OAC
Nitrogen atom count	NAC
Hydrogen-bond donors	HBD
Hydrogen-bond acceptor	HBA
Rotatable bonds	RB
Stereocenter count	SC
Topological polar surface area	TPSA
Number of rings	NOR
Aromatic ring count	ARC
Ring system count	RSC
Size of largest ring	SLR
Van der Waals surface area	VSA
Amine	-
Alcohol	-
Ether	-
Aldehyde	-
Ketone	-
Carboxylic Acid	-
Ester	-
Amide	-

Table E.2: Chapter 2 Chemical space coordinates of the screened chemicals

Chemical Space	PC1 (28% variance)	PC2 (19% variance)	PC3 (16% variance)	Members	Population (%)
1	-	-	-	9	15.8
2	+	-	-	8	14.0
3	-	-	+	6	10.5
4	+	-	+	8	14.0
5	+	+	-	2	3.5
6	-	+	-	13	22.8
7	+	+	+	6	10.5
8	-	+	+	5	8.8
			Total:	57	100

Figure E.1: Chapter 2 Eigenvalues for principal component determination

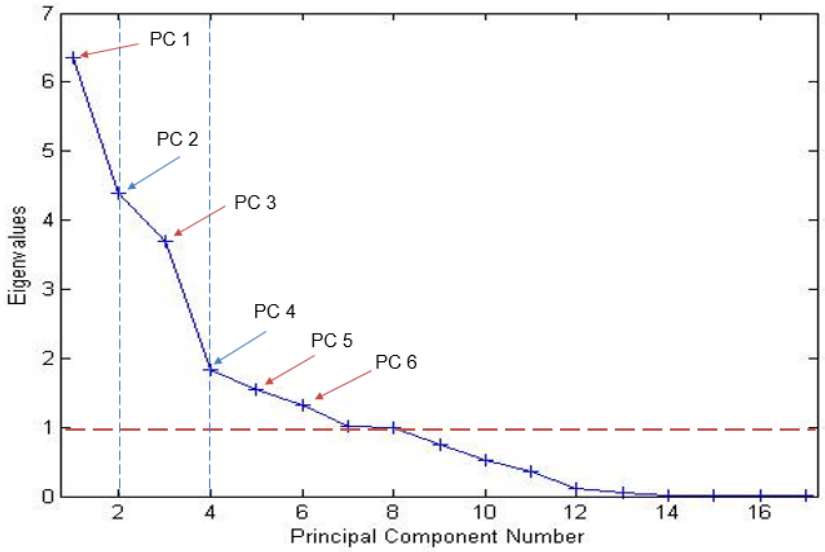

Figure E.1 – Two criteria to be considered in determining number of principal components are: 1) number of PCs with eigenvalues greater than one (red dashes and red arrows), 2) PCs located after bends or knees (blue dashes and blue arrow). Solo analysis resulted in six eigenvalues greater than one, thus six PCs (PCs 1 to 6) could be considered for the analysis. Two bends were also observed on the graph at PC2 and PC4, meaning that the PCs after the bends (PC3 and PC5) need to be

included in the analysis. These criteria led us to include three PCs (PCs 1-3) in our analysis.

APPENDIX F. CHEMICAL PANEL TABLES

Table F.1: Chapter 2 Table of chemicals in the chemical panel

Table F.1 Table of chemicals in the chemical panel

Chemical	CAS #	Supplier	Catalog #
(-)-limonene	5989-54-8	TCI Chemicals	L0132
(+)-carvone	6485-40-1	TCI Chemicals	C0703
(+)-menthol	2216-51-5	TCI Chemicals	M0826
1,2,3-trimethylbenzene	526-73-8	TCI Chemicals	T0468
1-butanol	71-36-3	TCI Chemicals	B0944
1-butoxy-2-propanol	5131-66-8	TCI Chemicals	B0864
1-dodecanol	112-53-8	Alfa Aesar	A12228
1-octanol	111-87-5	TCI Chemicals	O0036
1-propanol	71-23-8	TCI Chemicals	P0491
2,3-heptadione	96-04-8	TCI Chemicals	H0422
2,5-dimethylpyrazine	123-32-0	TCI Chemicals	D1526
2-methyl-1-butanol	137-32-6	TCI Chemicals	M0175
3-hexanone	589-38-8	TCI Chemicals	H0115
3-pentanol	584-02-1	TCI Chemicals	P0057
acetic acid	64-19-7	VWR	BDH3092
alpha-pinene	80-56-8	Sigma-Aldrich	305715
anisaldehyde	123-11-5	TCI Chemicals	A1674
beta-pinene	127-91-3	Sigma-Aldrich	CRM40433
butyric acid	107-92-6	TCI Chemicals	B0754
coumarin	91-64-5	TCI Chemicals	C0395
ethyl decanoate	110-38-3	TCI Chemicals	D0022
ethyl laurate	106-33-2	TCI Chemicals	L0013
ethyl myristate	124-06-1	TCI Chemicals	M0479
ethyl octanoate	106-32-1	TCI Chemicals	O0030
ethyl palmitate	628-97-7	Alfa Aesar	A15694
ethyl stearate	111-61-5	TCI Chemicals	S0079

eugenol	97-53-0	Alfa Aesar	A14332
farnesol	4602-84-0	TCI Chemicals	T0608
geraniol	106-24-1	TCI Chemicals	G0027
heptanoic acid	111-14-8	Alfa Aesar	A17704
hexanoic acid	142-62-1	Alfa Aesar	A13789
isovaleric acid	503-74-2	TCI Chemicals	M0182
lactic acid	50-21-5	Alfa Aesar	36415
lauric acid	143-07-7	TCI Chemicals	L0011
lilial	80-54-6	TCI Chemicals	B1145
linalool	78-70-6	TCI Chemicals	L0048
lyral	31906-04-4	Sigma-Aldrich	95594
melatonin	73-31-4	Sigma-Aldrich	M5250
methyl decanoate	110-42-9	Alfa Aesar	A15658
methyl laurate	111-82-0	Alfa Aesar	A12492
methyl myristate	124-10-7	Alfa Aesar	A10257
methyl octanoate	111-11-5	Alfa Aesar	A10991
methyl palmitate	112-39-0	TCI Chemicals	S0311
methyl stearate	112-61-8	TCI Chemicals	S0312
m-xylene	108-38-3	Sigma-Aldrich	95670
myristic acid	544-63-8	TCI Chemicals	M0476
nonanal	124-19-6	TCI Chemicals	N0296
octanal	124-13-0	Alfa Aesar	A10901
octanoic acid	124-07-2	Alfa Aesar	A11149
palmitic acid	57-10-3	TCI Chemicals	P0002
p-cymene	99-87-6	TCI Chemicals	C0513
pentadecanoic acid	1002-84-2	Alfa Aesar	A14664
propionic acid	79-09-4	TCI Chemicals	P05000
sotolon	28664-35-9	TCI Chemicals	D1851
stearic acid	57-11-4	TCI Chemicals	S0163
undecanal	112-44-7	TCI Chemicals	U0009
vanillin	121-33-5	Alfa Aesar	A11169

APPENDIX G. DNA SEQUENCES

G.1: Chapter 2 Sequences

G.1.1 Homo sapiens OR2A42 (UniprotKB - Q8NGT9)

ATGGGTGAAAATCAAACTATGGTTACCGAATTTCTACTACTAGGTTTCCTACT
AGGTCCAAGAATCCAAATGCTACTATTCGGTTTGTTCTCCTTGTTCTACATTTT
CACCTTGTTGGGTAACGGTGCTATTTTGGGTTTGATTTCCTTGGATTCCAGATT
GCATACCCCAATGTATTTCTTCCTAAGTCATTTGGCTGTTGTTGATATTGCTTA
CACCAGAAACACCGTTCCACAAATGTTGGCTAACTTGTTGCACCCTGCCAAG
CCTATTAGCTTCGCCGGATGCATGACCCAAACCTTCTTATGTTTATCCTTCGGT
CATTCCGAATGTTTGTTGTTGGTTTTGATGTCTTACGATAGATACGTTGCTATT
TGTCATCCATTGAGATACTCCGTTATTATGACTTGGAGAGTTTGTATTACTTTG
GCTGTTACTTCCTGGACTTGTGGTTCCTTGTTGGCTTTGGCTCATGTTGTTTTG
ATTTTGAGATTGCCATTCTCCGGCCCACATGAAATTAACCATTTCTTCTGTGA
AATTTTGTCCGTTTTGAGATTGGCTTGTGCTGATACTTGGTTGAACCAAGTTG
TTATTTTCGCTGCTTGTGTCTTCTTCCTAGTTGGTCCACCATCTTTGGTTTTGGT
TTCTTACTCTCATATTTTGGCTGCTATTTTGAGAATACAATCTGGCGAAGGCA
GAAGAAAGGCTTTCTCTACTTGTTCTTCTCATTTGTGTGTTGTTGGTTTGTTCT
TCGGTTCTGCTATTATTATGTACATGGCTCCAAAGTCTAGACATCCAGAAGAA
CAACAAAAGGTTTTCTTCTTGTTCTACTCTTTCTTCAACCCAACTTTGAACCCA
TTGATTTACTCTTTGAGAAACGGCGAAGTTAAGGGTGCTTTGAGGAGGGCTTT
GGGCAAGGAATCTCATTCTTAA

G.1.2 Homo sapiens OR2W3 (UniprotKB - Q7Z3T1)

ATGGATGGTACTAACGGTTCCACCCAAACCCACTTTATTCTATTGGGTTTCTC
CGATAGACCACATTTGGAAAGAATTTTGTTCGTTGTTATTTTGATTGCTTACTT
GTTGACCTTGGTTGGTAACACCACCATTATTTTGGTTTCCAGACTAGACCCAC
ATTTGCATACCCCAATGTACTTCTTCTTGGCTCATTTGTCCTTCTTGGATTTGT
CTTTCACCACCTCCTCCATTCCACAATTGTTGTACAACTTGAACGGTTGTGAT
AAGACCATTTCCTACATGGGTTGTGCTATTCAATTGTTCTTGTTCTTGGGTTTG
GGTGGTGTTGAATGTTTGTTGTTGGCTGTTATGGCTTACGATAGATGTGTTGC
TATTTGTAAGCCATTGCATTACATGGTTATTATGAACCCAAGATTGTGTAGAG
GTTTGGTTTCTGTTACCTGGGGTTGTGGTGTCGCTAACTCTTTGGCTATGTCTC
CAGTCACTCTACGTCTACCAAGATGTGGCCATCATGAAGTTGATCATTTCTTG
AGAGAAATGCCAGCTTTGATTAGAATGGCTTGTGTTTCTACTGTTGCTATTGA
AGGTACTGTTTTCGTTTTGGCTGTTGGTGTTGTTCTAAGTCCATTGGTTTTCAT
TTTGTTGTCTTACTCTTACATTGTTAGAGCTGTTCTACAAATTAGATCTGCTTC
TGGTAGACAAAAGGCTTTCGGTACTTGTGGTTCTCATTTGACTGTTGTTTCTTT
GTTCTACGGTAACATTATTTACATGTACATGCAACCAGGTGCTTCTTCTTCTC
AAGATCAAGGTATGTTTCTAATGCTATTCTACAACATTGTTACTCCATTGTTG
AACCCCCTAATCTACACTCTACGTAACCGTGAAGTCAAGGGTGCTTTGGGTA
GATTGTTGTTGGGTAAGAGAGAACTAGGCAAGGAATAA

G.1.3 Homo sapiens OR2T4 (UniprotKB - Q8NH00)

ATGGGTCTGTTCAGACAGAGCAAGCACCCAATGGCTAACATTACCTGGATGG
CTAACCATACCGGTTGGAGCGATTTCATTCTGCTGGGTCTGTTCAGACAATCC

AAGCATCCAGCTCTGCTGTGTGTCGTCATTTTCGTTGTTTTCTTGATGGCTTTG

TCCGGTAACGCTGTCCTAATCTTGTTGATTCATTGTGATGCTCATTTGCATACC

CCAATGTACTTCTTCATTTCCCAATTGTCCTTGATGGATATGGCTTACATTTCC

GTTACCGTTCCAAAGATGTTGTTGGATCAAGTTATGGGTGTTAACAAGATTTC

CGCTCCAGAATGTGGTATGCAAATGTTCTTCTACGTTACCTTGGCTGGTTCCG

AGTTCTTTCTGTTGGCTACTATGGCTTACGATAGATACGTTGCTATTTGTCATC

CATTGAGATACCCAGTTTTGATGAACCATAGAGTTTGTTTGTTCTTGTCCTCC

GGTTGTTGGTTCTTGGGTTCCGTCGATGGTTTTACCTTCACCCCAATTACCATG

ACCTTCCCATTCAGAGGTTCCAGAGAAATTCACCACTTCTTCTGTGAAGTTCC

AGCTGTTTTGAACTTGTCCTGTTCCGATACTTCCTTGTACGAAATTTTCATGTA

CTTGTGTTGTGTTTTGATGTTGTTGATTCCAGTTGTTATTATTTCTTCTTCTTAC

TTGTTGATTTTGTTGACTATTCATGGTATGAACTCTGCTGAAGGTAGAAAGAA

GGCTTTCGCTACTTGTTCTTCTCATTTGACTGTTGTTATTTTGTTCTACGGTGCT

GCTATTTACACTTACATGTTGCCATCTTCTTACCATACTCCAGAGAAGGATAT

GATGGTTTCTGTTTTCTACACTATTTTGACTCCAGTTGTTAACCCATTGATTTA

CTCTTTGCGTAACAAGGATGTTATGGGTGCTCTAAAGAAGATGTTGACTGTTG

AACCAGCCTTCCAAAAGGCAATGGAATAA

G.1.4 Homo sapiens OR51B5 (UniprotKB - Q9H339)

ATGTCCAGCTCCGGCTCCTCCCATCCATTCTTGTTGACCGGCTTCCCAGGTTT

GGAAGAAGCTCATCATTGGATTTCCGTTTTCTTCTTGTTCATGTACATTTCCAT

TTTGTTCGGTAACGGTACCCTACTACTACTAATTAAGGAAGATCATAACTTGC

ATGAGCCTATGTATTTCTTCTTGGCTATGTTGGCTGCTACCGATTTGGGTTTGG

CTTTGACCACCATGCCAACCGTTTTGGGTGTTTTGTGGTTGGATCATAGAGAG

ATTGGTTCTGCCGCCTGTTTCTCCCAAGCCTACTTCATTCATTCCTTGTCCTTC

TTGGAATCCGGCATTTTGTTGGCTATGGCTTACGATCGTTTCATTGCTATTTGT

AACCCATTGAGATACACCTCCGTTCTAACTAATACTCGTGTTGTTAAGATTGG

CTTGGGTGTTTTGATGAGAGGTTTCGTTTCCGTTGTTCCACCAATTAGACCATT

GTACTTCTTCTTGTACTGTCATTCCCATGTTTTGTCCCATGCTTTCTGTTTGCAT

CAAGATGTTATTAAGTTGGCTTGTGCTGATACTACTTTCAACAGATTGTACCC

AGCTGTTTTGGTTGTTTTCATTTTCGTTTTGGATTACTTGATCATCTTCATTAGT

TACGTCCTAATTTTGAAGACTGTTTTGTCCATTGCTTCCAGAGAAGAAAGAGC

TAAGGCTTTGATTACTTGTGTTAGTCATATTTGTTGTGTTTTGGTTTTCTACGT

TACTGTTATTGGTCTAAGTTTGATTCATCGTTTCGGAAAACAGGTCCCACATA

TTGTTCATTTGATTATGTCCTACGCTTACTTCCTATTCCCTCCACTAATGAACC

CAATTACTTACTCCGTTAAGACCAAGCAAATTCAAAACGCTATTTTGCATTTG

TTCACCACCCATAGAATTGGTACCTAA

G.1.5 Homo sapiens OR2L13 (UniprotKB - Q8N349)

ATGGAGAAGTGGAATCATACATCAAACGACTTCATCCTTCTGGGTCTGTTACC

TCCTAATCAGACCGGTATCTTCCTGCTCTGTCTGATCATTCTGATCTTCTTCCT

GGCAAGCGTTGGTAATAGCGCAATGATTCATCTGATTCACGTTGATCCGCGTC

TGCATACCCCGATGTATTTCCTGCTGAGCCAGCTGAGCCTGATGGATCTGATG

TATATTAGTACCACCGTTCCGAAGATGGCATATAATTTCTTAAGTGGTCAGAA

GGGCATCAGTTTCTTAGGTTGTGGTGTTCAAAGTTTCTTCTTCTTAACAATGG

CCTGTAGTGAAGGTCTGTTACTCACCTCAATGGCCTACGATCGTTATTTAGCA

ATTTGTCATTCATTATATTATCCTATTAGAATGTCAAAGATGATGTGCGTTAA

GATGATTGGTGGTTCTTGGACATTAGGTTCAATTAACAGTTTAGCACATACAG

TGTTCGCATTACATATTCCGTATTGTAGAAGTAGAGCAATTGATCATTTCTTC

TGTGACGTTCCTGCTATGTTATTATTAGCTTGTACCGATACCTGGGTTTACGA

ATATATGGTGTTCGTTAGTACCAGTCTGTTCTTGCTGTTCCCATTCATTGGTAT

TACCAGTAGTTGTGGTCGTGTTCTGTTCGCAGTTTATCACATGCATAGCAAGG

AAGGCAGAAAGAAGGCATTCACTACTATTAGCACCCATCTGACCGTTGTTAT

CTTCTATTACGCACCTTTCGTTTACACCTATCTGCGTCCTCGTAACCTGCGTAG

TCCGGCAGAAGATAAGATTCTGGCAGTGTTCTACACCATTCTGACCCCGATGC

TGAACCCGATTATTTATAGCCTGCGTAATAAAGAAGTTCTGGGTGCAATGCGT

CGTGTGTTCGGTATCTTCAGCTTCCTGAAGGAATAA

G.1.6 Homo sapiens OR10S1 (UniprotKB - Q8NGN2)

ATGACCAGTCGTAGTGTTTGTGAAAAGATGACCATGACCACCGAAAATCCGA

ATCAGACAGTTGTTAGCCATTTCTTCTTAGAAGGTTTACGTTATACAGCAAAG

CATAGTAGTTTATTCTTCTTATTATTCTTATTAATTTACAGTATTACAGTTGCA

GGTAACTTATTAATTTTATTAACCGTTGGTAGTGATAGCCATTTAAGTTTACC

GATGTATCATTTCTTAGGTCATTTAAGTTTCTTAGACGCCTGTTTAAGTACCGT

TACAGTTCCGAAGGTTATGGCAGGTTTATTAACCTTAGACGGTAAAGTTATTA

GTTTCGAAGGTTGTGCAGTTCAGTTATATTGTTTCCACTTCTTAGCAAGTACA

GAGTGTTTCTTATATACAGTTATGGCATACGATCGTTATTTAGCAATCTGTCA

GCCGTTACATTATCCGGTTGCCATGAATCGTCGTATGTGTGCCGAAATGGCAG

GTATTACCTGGGCAATTGGTGCAACCCACGCAGCAATTCATACCAGCCTGAC

CTTTCGTCTGCTGTATTGTGGTCCGTGTCATATTGCATATTTCTTCTGTGATAT

TCCGCCGGTTCTGAAACTGGCCTGTACCGATACCACCATTAACGAACTGGTTA

TGTTAGCAAGTATTGGTATTGTTGCAGCAGGTTGTCTGATTCTGATTGTTATT

AGTTATATTTTCATTGTTGCAGCAGTTTTACGTATTCGTACAGCACAGGGTCG

TCAGCGTGCATTCAGTCCGTGTACAGCACAATTAACAGGTGTTTTATTATATT

ACGTTCCGCCGGTTTGCATTTATTTACAACCGCGTAGCAGTGAAGCAGGTGCC

GGTGCACCGGCAGTTTTCTATACCATTGTTACACCGATGCTGAATCCGTTTAT

TTATACATTACGTAATAAGGAGGTTAAACACGCATTACAACGTTTACTGTGTA

GTAGTTTTCGTGAAAGCACCGCAGGTAGCCCGCCGCCGTAA

G.1.7 Homo sapiens OR2A7 (UniprotKB - Q96R45)

ATGGGTGACAACATTACCAGCATTACCGAATTTCTGCTGCTGGGTTTTCCTGT

TGGTCCACGCATCCAGATGCTGCTGTTTGGTCTGTTCAGCCTGTTCTACGTATT

TACACTGCTGGGTAACGGTACCATTTTAGGTTTAATTAGCTTAGATAGCCGTT

TACACGCACCTATGTACTTCTTCTTAAGCCATTTAGCAGTTGTTGATATTGCAT

ACGCTTGCAACACAGTTCCTCGTATGTTGGTTAACTTATTACATCCGGCTAAA

CCTATTAGCTTCGCAGGTCGTATGATGCAAACCTTCTTATTTAGCACCTTCGC

AGTTACCGAGTGTTTATTATTAGTTGTTATGAGCTACGATTTATACGTTGCAA

TTTGCCATCCATTACGTTACTTAGCAATTATGACCTGGCGTGTTTGTATTACCT

TAGCAGTTACAAGCTGGACAACAGGTGTTCTGTTAAGTCTGATTCACTTAGTT

TTATTATTACCTTTACCTTTCTGTCGTCCTCAAAAGATTTATCATTTCTTCTGT

GAAATTTTAGCAGTTTTAAAGTTAGCCTGTGCAGACACACACATTAACGAAA

ACATGGTTTTAGCAGGTGCAATTAGCGGTTTAGTTGGTCCTTTAAGTACCATT

GTTGTTAGCTACATGTGTATTCTGTGTGCAATTTTACAAATTCAAAGCCGTGA

AGTTCAGCGTAAGGCATTCTGCACCTGTTTCAGCCACCTGTGTGTTATTGGTT

TATTCTACGGTACAGCAATTATTATGTACGTTGGTCCGCGTTACGGTAACCCT

AAGGAACAAAAGAAGTACTTATTATTATTCCATAGCCTGTTTAACCCGATGCT

GAACCCGCTGATTTGTAGCCTGCGTAACAGCGAAGTTAAGAACACCCTGAAG
CGCGTTCTGGGTGTTGAACGTGCACTGTAA

G.2: Chapter 3 Sequence

G.2.1 Yeast-optimized NanoLuc sequence

ATGGTCTTCACCTTGGAAGATTTCGTCGGTGATTGGAGACAAACCGCTGGTTA
CAACTTGGATCAAGTCTTAGAGCAGGGTGGAGTCTCCTCCTTGTTTCAAAACT
TGGGAGTCTCCGTCACCCCAATTCAAAGAATTGTTTTGTCCGGTGAAAATGGT
TTGAAAATTGATATTCATGTCATTATTCCTTACGAAGGTTTGTCCGGTGATCA
AATGGGTCAAATTGAAAAGATTTTCAAGGTCGTCTACCCAGTTGATGATCATC
ATTTTAAAGTCATTTTGCATTACGGTACCTTGGTCATTGATGGTGTCACCCCA
AACATGATTGATTACTTTGGTAGACCTTACGAAGGTATTGCTGTCTTTGATGG
TAAGAAGATTACCGTTACTGGTACCTTGTGGAATGGTAACAAGATTATTGAT
GAAAGATTGATTAACCCAGATGGTTCCTTGTTGTTTAGAGTCACTATTAATGG
TGTCACTGGTTGGAGATTGTGTGAAAGAATTTTGGCTTAA